



Research paper

Community detection via deep motif-regularized asymmetric nonnegative matrix factorization

Hazhir Sohrabi ^a, Amjad Seyedi ^b, Shahrokh Esmaili ^a,* , Parham Moradi ^c

^a Department of Applied Mathematics, University of Kurdistan, Sanandaj, Iran

^b Department of Mathematics and Operational Research, University of Mons, Mons, Belgium

^c College of Arts, Business, Law Education, and Information Technology, Victoria University, Melbourne, Australia

ARTICLE INFO

Keywords:

Community detection
Nonnegative matrix factorization
Diagonal dominance
Structural proximity
Interpretable clustering

ABSTRACT

Community detection in complex networks faces persistent challenges in modeling hierarchical organization, higher-order structures, and directional interactions. Deep Nonnegative Matrix Factorization (Deep NMF) models effectively reveal network hierarchical structures, supporting community detection. However, they often focus on first-order connections, overlooking higher-order patterns like network motifs, recurring subgraphs that capture complex interactions. By bridging the concept of motifs with deep representation learning, we establish new possibilities for analyzing modern network paradigms. We propose Deep Motif-Regularized Asymmetric NMF (DMRA-NMF), a deep model that synergizes multi-scale factorization with motif-aware regularization to address these limitations. Our method constructs hybrid similarity matrices integrating edge connectivity and statistically significant motifs, while employing asymmetric factorization to preserve directional network flows. A novel diagonal dominance constraint sharpens community boundaries by suppressing inter-cluster noise without sacrificing intra-community cohesion. The unified optimization strategy balances local neighborhood preservation with global structural coherence, offering interpretable community structures while maintaining computational efficiency. Extensive experiments on biological, social, and technological networks demonstrate the framework's superiority in detecting nested communities and resolving boundary ambiguities compared to eleven state-of-the-art methods.

1. Introduction

The analysis of complex networks has become a cornerstone of modern data science, providing critical insights into systems as diverse as social interactions, biological processes, and technological infrastructures (Benson et al., 2016). At the heart of this analysis lies community detection—the identification of densely connected node groups that often correspond to functional modules or organizational units. These communities reveal latent patterns in social circles (Clauset et al., 2004), protein interaction clusters (Hastings, 2006), and multi-layer network dynamics (Gao et al., 2023). Despite decades of research, existing methods face persistent challenges in addressing three intrinsic properties of real-world networks: hierarchical organization spanning micro-to-macro scales, the prevalence of higher-order structural motifs beyond dyadic edges, and directional asymmetries in interaction patterns that defy symmetric modeling assumptions.

Nonnegative Matrix Factorization (NMF) (Lee and Seung, 1999; Gillis, 2020) is a widely used technique for dimensionality reduction that breaks down a high-dimensional nonnegative data matrix

into two lower-dimensional nonnegative matrices. The non-negativity constraints promote a parts-based representation, aligning with the intuitive idea of combining parts to form a whole. Because of this interpretability, NMF has found broad applications across various domains, community detection (Ghodsi et al., 2025), link prediction (Mahmoodi et al., 2024), recommendation (Gouvert et al., 2020), hyperspectral unmixing (Feng et al., 2022), and pattern recognition (Barkhoda et al., 2026). Basic NMF provides node representations but is not inherently designed for graph clustering. To address this, Kuang et al. (2012) proposed symmetric NMF (SymNMF), which factorizes a symmetric affinity matrix $A \approx HH^T$, where H indicates soft cluster membership. By preserving the graph's structure, SymNMF enables more effective and interpretable graph clustering than standard NMF. This traditional approach, while successful in simple networks (He et al., 2022), exhibits critical shortcomings when applied to modern network paradigms. Shallow factorization architectures like SymNMF (Kuang et al., 2012) and Asymmetric NMF (ANMF) (Wang et al., 2011) fail to capture nested community hierarchies observed in organizational systems. The

* Corresponding author.

E-mail addresses: sh.esmaeili@uok.ac.ir (S. Esmaili), parham.moradi@vu.edu.au (P. Moradi).

reliance on adjacency matrices alone ignores statistically significant subgraphs (motifs) that encode functional relationships, while symmetric factorization proves inadequate for directed networks like citation graphs or metabolic pathways (Malliaros and Vazirgiannis, 2013).

To overcome these limitations, Deep NMF methods extend traditional NMF into multi-layer architectures, enabling the learning of hierarchical and more abstract representations—analogueous to deep neural networks (Trigeorgis et al., 2014). While deep NMFs have been applied in network analysis, they primarily serve as multi-layer node representation models and lack intrinsic clustering capabilities (De Handschutter et al., 2021). To address the limitations of earlier models, Deep Symmetric NMF (DSNMF) (De Handschutter et al., 2023) incorporates multi-layer symmetric regularization for hierarchical clustering in undirected graphs. More recently, Deep Asymmetric NMF (DAs-NMF) (Hajiveisheh et al., 2024) extends this approach to directed and undirected networks by directly performing graph clustering through multi-layer factorization. It captures edge directionality while jointly learning node representations and graph-level summaries, eliminating the need for separate clustering steps.

Another recent effort involves incorporating motif-based similarities (Shang et al., 2022; Wu et al., 2024) and graph regularization (Deng et al., 2021) have yielded incremental improvements but lack a unified framework for hierarchical, direction-aware analysis. A key challenge in community detection is mitigating unintended connections between subgraphs. This ensures nodes with robust intra-group ties remain clustered within the same community, avoiding misclassification and enhancing clustering precision (Li et al., 2024b). Implementing diagonal dominance constraints on the interaction matrix can resolve this issue. By ensuring that diagonal entries (representing intra-community interactions) are substantially larger than off-diagonal entries (inter-community interactions), this approach minimizes artificial correlations between distinct subgraphs. Such constraints sharpen community boundaries, reduce noise in inter-cluster relationships, and improve result interpretability (Zhang et al., 2020).

A notable limitation in existing community detection methods is their shallow architectural design, which often restricts them to a single-level understanding of network structures. Such flat factorizations are inherently incapable of modeling the multi-resolution nature of real-world networks, where communities are nested and hierarchically organized across scales. Recent works have hinted at the potential of deep matrix factorization techniques to overcome this limitation by enabling layer-wise abstraction of community structures (Zhang and Zhou, 2020; Wang and Zhang, 2023). However, these efforts remain fragmented and lack a comprehensive framework that systematically incorporates depth while addressing other structural nuances like motif prevalence and directional interactions. This research gap calls for the development of a deep factorization-based model that can reveal hierarchical community patterns more effectively.

To bridge these gaps, we propose a framework that synergizes four pivotal advancements in network science: Deep Hierarchical Factorization, a deep architecture employing sequential factorizations to uncover nested communities at varying resolutions; Mixed-Order Similarity Integration, adaptive fusion of first-order adjacency relations with motif-derived higher-order proximities; Directional Asymmetry Preservation, mechanisms to maintain interaction directionality through asymmetric matrix operations; Topology-Conscious Regularization, a regularization term preserving local node neighborhoods while suppressing inter-community overlaps; and Diagonal Dominance Block Matrix, by employing diagonal dominance constraints, we enhance intra-community interactions, minimize inter-community noise, and sharpen clustering accuracy in our method. The proposed framework introduces three fundamental advances over state-of-the-art approaches. The first innovation involves motif-aware representation learning through the construction of hybrid similarity matrices that combine traditional edge-based connections (\mathbf{A}) with motif adjacency matrices (\mathbf{B}_{M_i}). This integration addresses the sparsity limitations of conventional adjacency matrices

while effectively capturing the functional dependencies encoded in triadic motifs, which proves particularly crucial for directed networks where specific motif patterns like feed-forward loops play governing roles in system behavior (Milo et al., 2002; Benson et al., 2016).

The second advancement comes from the deep asymmetric architecture, which fundamentally differs from shallow NMF variants. This deep factorization approach enables progressive refinement of community memberships across hierarchical levels, with each layer's factorization encoding increasingly abstract community relationships. This architecture successfully mimics the multi-scale organization observed in real networks.

The framework incorporates two complementary regularization strategies: graph Laplacian regularization preserves local neighborhood structures by minimizing discrepancies between connected nodes in the embedding space, ensuring that nodes with similar connectivity patterns remain proximate in the learned representation; while diagonal dominance constraints promote clearer community boundaries through targeted regularization of off-diagonal elements in the interaction matrix, specifically focusing on sparsifying inter-community connections without making global assumptions about the network structure. The experimental results highlight the framework's robustness across various network domains and its ability to address the fundamental challenges of hierarchical organization, higher-order structure incorporation, and directional relationship preservation that limit current methods.

Community detection in complex networks remains a fundamental task in network science, yet it continues to face challenges due to the heterogeneity of interaction patterns and the presence of higher-order structures. While existing NMF-based approaches have shown promise in uncovering communities, they often rely on shallow architectures and symmetric assumptions that limit their capacity to model directional interactions and capture deep structural dependencies. Moreover, most methods overlook the influence of network motifs—small recurring subgraph patterns—that play a crucial role in shaping the global organization of real-world networks. To address these limitations, we propose a deep motif-regularized asymmetric nonnegative matrix factorization (DMRA-NMF) model that integrates hierarchical factorization with motif-aware regularization and asymmetric structure modeling. Our goal is to more accurately extract communities by leveraging motif patterns, preserving local structure, and enforcing diagonal dominance in the latent space. This combination enables DMRA-NMF to capture both low-level and high-order proximities, leading to more robust and interpretable community structures.

The proposed framework introduces several key innovations in the field of community detection:

- It employs Symmetric Nonnegative Matrix Tri-Factorization (SNMTF) with an explicit diagonalization constraint on the middle matrix to encourage more distinct and interpretable community memberships.
- It incorporates structural proximity through a similarity-based regularization term that preserves local structure in the network, improving robustness to noise and sparsity.
- The method jointly optimizes both diagonal enhancement and local structure preservation, integrating these factors directly into the factorization objective function.
- The framework improves the interpretability and accuracy of detected communities while maintaining the symmetry and sparsity of the original adjacency matrix.
- It reduces model complexity by avoiding multi-stage learning schemes and instead formulates a unified optimization problem.
- Moreover, the method leverages network motifs as higher-order structural patterns to capture more nuanced community structures beyond simple pairwise connections.

The remainder of this paper is organized as follows: Section 2 formalizes the community detection problem and reviews NMF fundamentals. Section 3 details the proposed methodology, including motif adjacency construction and optimization algorithms. Sections Section 4 present experimental results and comparative analyses. Section 5 discusses limitations and future directions, followed by concluding remarks.

2. Background

2.1. Graph and matrix notation

For an arbitrary matrix $\mathbf{A} \in \mathbb{R}^{n \times k}$, A_{ij} denotes the i, j entry of \mathbf{A} , while $A_{i \cdot}$ and $A_{\cdot j}$ represent the i th row and j th column of \mathbf{A} , respectively. Let $G = (V, E)$ represent a graph with vertex set $V = \{v_1, v_2, \dots, v_n\}$ ($|V| = n$) and edge set $E \subseteq V \times V$. The adjacency matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ has elements $A_{ij} = 1$ if $(v_i, v_j) \in E$ and 0 otherwise. In this work, we employ several fundamental operations: Frobenius norm $\|\cdot\|_F$ for reconstruction error minimization, Hadamard product \circ for element-wise matrix operations, all-ones matrix $\mathbf{1}$ and identity matrix \mathbf{I} for regularization, and trace function $\text{Tr}(\cdot)$ for matrix norm computations.

2.2. Related work

NMF and its variants have emerged as powerful tools for community detection, offering interpretable low-rank representations of network structures. The standard NMF factorizes a nonnegative matrix $\mathbf{X} \in \mathbb{R}_+^{m \times n}$ into $\mathbf{X} \approx \mathbf{W}\mathbf{H}$, where \mathbf{W} encodes features and \mathbf{H} represents sample embeddings (Lee and Seung, 1999; Wang et al., 2008). For symmetric adjacency matrices, SymNMF enforces $\mathbf{A} \approx \mathbf{H}\mathbf{H}^T$ to preserve network symmetry (Kuang et al., 2012), while ANMF addresses directed graphs through tri-factorization $\mathbf{A} \approx \mathbf{H}\mathbf{W}\mathbf{H}^T$, where \mathbf{W} captures directional inter-cluster relationships (Wang et al., 2011). Recent extensions, such as Liu et al. (2023), propose a symmetry and graph bi-regularized NMF (B-NMF) to enhance representation learning in large-scale undirected networks by leveraging multiple latent factor matrices under equality constraints, thereby preserving intrinsic symmetry while incorporating local geometry via graph regularization.

Recent advances extend NMF to deep architectures for hierarchical community detection. By factorizing $\mathbf{X} \approx \mathbf{W}_1\mathbf{W}_2 \dots \mathbf{W}_p\mathbf{H}_p$, deep NMF (DNMF) captures multi-scale patterns through layer-wise transformations \mathbf{W}_i , where implicit representations $\mathbf{H}_{i-1} \approx \mathbf{W}_i\mathbf{H}_i$ enable progressive feature abstraction (Pei et al., 2019; Wang and Zhang, 2023). Building on this, Li et al. (2024a) introduces Contrastive Deep NMF (CDNMF), which employs contrastive learning to unify network topology and node attributes as dual views. While CDNMF advances deep NMF by integrating multi-modal data, it overlooks higher-order structural motifs critical in asymmetric networks. In contrast, Hajiveiseh et al. (2024) proposes Deep Asymmetric NMF (DAsNMF) for directed graphs, incorporating asymmetric cosine and PageRank-based similarities to preserve local/global structures. However, DASNMF focuses on pseudo-hierarchical clustering without explicit motif integration—a gap addressed in our work.

The integration of motif analysis with NMF has shown promise in biological networks. Hutchins et al. (2008) demonstrates how NMF can simultaneously characterize both positioning and sequence content of regulatory motifs, while Kim et al. (2011b) introduces biological network motif detection using NMF-based approaches. Aadhithya et al. (2023) provides a comparative analysis of motif detection methods, highlighting the trade-offs between accuracy and computational efficiency in NMF-based approaches. For protein sequence analysis, Kim et al. (2011a) and Aghdam and Zanjani (2021) utilize sparse NMF (SNMF) for motif discovery, demonstrating improved clustering of protein structures. In directed networks, recent work has focused on

asymmetric NMF variants. Yu et al. (2024) proposes a robust asymmetric NMF using Cauchy distributions to handle non-Gaussian errors, while Tosyali et al. (2019) introduces regularized asymmetric NMF (RANMF) that incorporates pairwise node similarity. Aghdam and Zanjani (2021) applies regularized asymmetric NMF to text clustering, demonstrating improved discrimination through distance-based regularization.

While effective, shallow and Deep NMF variants often neglect higher-order structural patterns. Motif-aware approaches like Mixed-Order NMF (MONMF) (Bu et al., 2025) address this by incorporating motif adjacency matrices $\mathbf{B}_{\mathcal{M}}$, computed as $\mathbf{B}_{\mathcal{M}} = \sum_c (\mathbf{D}_c \cdot \mathbf{E}_c) \circ \mathbf{F}_c + \text{transposed terms}$, where \mathbf{D}_c , \mathbf{E}_c , and \mathbf{F}_c encode bidirectional, unidirectional, and missing edges (Benson et al., 2016). Complementary to motif-based methods, Wu et al. (2023) introduces Graph Embedding via Motif-aware Feature Propagation (GEMFP), blending first-order adjacency with higher-order motifs through iterative influence aggregation. Mixed-order feature matrices \mathbf{B}_{MX} further balance these via $\mathbf{B}_{MX} = (1 - \lambda)\mathbf{B}_{\mathcal{M}} + \lambda\mathbf{A}$, but lack deep factorization capabilities (Shang et al., 2022).

To preserve intrinsic geometric structures, Graph regularized NMF (GNMF) introduces manifold constraints through Laplacian regularization. The graph Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{S}$, where \mathbf{S} is the similarity matrix, enforces smoothness in embeddings via $\text{Tr}(\mathbf{U}^T \mathbf{L} \mathbf{U})$ (Deng et al., 2010) and \mathbf{D} is diagonal degree matrix of \mathbf{S} ($D_{ii} = \sum_j S_{ij}$). Tang et al. (2017) unifies global node importance and local node similarities in an NMF framework. Similarly, Mohammadi et al. (2024) proposes Diverse Joint NMTF (Div-JNMTF) for attributed graphs, using HSIC-based diversity regularization to decouple topological and non-topological representations. Recent work by Luo et al. (2024) extends this paradigm by integrating node content through dual encoder–decoder architectures (DEDNMF and DDNMF), imposing sparsity constraints to balance topology and content adaptively. Similarly, Zhao et al. (2022) proposes Adaptive NMF (AdNMF) with flexible topology-content trade-offs using coupled encoder–decoder frameworks. While these methods unify structural and attribute data, they lack explicit modeling of directional relationships and multi-layer motif integration. As summarized in Table 1, our proposed method uniquely integrates all key features – deep architecture, motif awareness, and adaptive diagonal constraints – outperforming existing approaches in comprehensive capability.

Our work bridges these gaps by proposing an AsNMF model that jointly leverages motifs and Laplacian regularization. Unlike Hajiveiseh et al. (2024), we incorporate motif adjacency matrices to encode higher-order interactions while employing asymmetric tri-factorization to preserve directional relationships. The Laplacian constraint, inspired by Tang et al. (2017), ensures structural coherence in the deep latent space, advancing beyond shallow implementations in Mohammadi et al. (2024). This unified approach surpasses hybrid methods like Wu et al. (2023) by enabling multi-layer factorization of mixed-order structures, while HSIC-inspired regularization (Mohammadi et al., 2024) enhances interpretability through disentangled representations.

This work attempts to bridge the gaps identified in the aforementioned methods by proposing a deep asymmetric NMF framework, termed DMRA-NMF, which jointly leverages motif-aware regularization and Laplacian structural preservation to enhance representation learning and clustering performance. Unlike DASNMF (Hajiveiseh et al., 2024), which captures directionality but neglects higher-order structural patterns, our model integrates motif adjacency matrices to encode triadic and feed-forward dependencies that are essential in directed networks. Compared with MONMF (Tang et al., 2017), which applies motif-based regularization within a shallow structure, DMRA-NMF embeds the motif information into a deep hierarchical tri-factorization, enabling multi-level community abstraction. Furthermore, unlike GEMFP (Wu et al., 2023), which fuses edge- and motif-level proximities without adaptive control, our mixed-order similarity integration adaptively balances first- and higher-order relations.

Table 1

Comprehensive comparison of representative NMF-based methods in terms of key model characteristics. A checkmark (✓) indicates that the corresponding feature is supported by the method, while a cross (✗) denotes its absence. The proposed model integrates all the listed properties, including deep factorization, local geometry preservation, tri-factor structure, adaptive diagonal constraint, community correlation, hybrid graph similarity, and motif theory integration.

Method	Deep	Local	Tri-Factor	Adaptive Diag.	Correlation	Hybrid	Motif
NMF	✗	✗	✗	✗	✗	✗	✗
SymNMF	✗	✗	✗	✗	✗	✗	✗
ANMF	✗	✗	✓	✗	✗	✗	✗
B-NMF	✗	✓	✗	✗	✗	✗	✗
DNMF	✓	✗	✗	✗	✗	✗	✗
CDNMF	✓	✗	✗	✗	✗	✗	✗
DEDNMF	✗	✓	✗	✗	✗	✗	✓
DDNMF	✗	✓	✗	✗	✗	✗	✓
DASNMF	✓	✓	✓	✗	✗	✓	✗
MONMF	✗	✗	✗	✗	✗	✓	✓
GEMFP	✗	✗	✗	✗	✗	✓	✓
GNMF	✗	✓	✗	✗	✗	✗	✗
Div-JNMTF	✗	✗	✓	✗	✗	✗	✗
Proposed	✓	✓	✓	✓	✓	✓	✓

Deep = Deep factorization Correlation = Community correlation

Local = Local geometry preservation Hybrid = Structure-based graph similarity

Tri-Factor = Three-matrix factorization Motif = Motif theory integration

Adaptive Diag. = Adaptive diagonal constraint.

The Laplacian constraint ensures structural coherence in the deep latent space, while the adaptive diagonal constraint enhances intra-community compactness and interpretability. Inspired by the HSIC-based disentanglement mechanism in [Mohammadi et al. \(2024\)](#), our framework further refines latent representations to achieve more interpretable and structurally consistent communities. Together, these mechanisms allow DMRA-NMF to preserve directional, hierarchical, and motif-based dependencies more effectively than existing asymmetric or motif-aware NMF variants.

3. Proposed method

3.1. Motivation

The analysis of modern complex networks reveals three persistent challenges that existing community detection methods fail to address simultaneously: hierarchical organization across multiple resolution levels, higher-order interactions beyond dyadic edges, and directional asymmetries in network flows. Current approaches remain fragmented—while SymNMF handles undirected networks ([Kuang et al., 2012](#)) and ANMF processes directed graphs ([Wang et al., 2011](#)), these shallow factorizations cannot capture multi-scale community structures. Moreover, motif-agnostic methods overlook statistically significant subgraphs that encode functional relationships ([Milo et al., 2002](#)), and symmetric formulations distort directional dependencies in critical networks like citation graphs and metabolic pathways. This methodological fragmentation motivates our integrated solution, inspired by the empirical observation of layered motif participation in hierarchical networks, where nodes engage in distinct motif types at different structural depths (e.g., feed-forward loops at macro scales versus transitive triangles locally). Our proposed framework addresses these limitations through three coordinated innovations: a multi-layer factorization that progressively resolves communities from coarse to fine granularity, motif-Laplacian co-regularization preserving both local neighborhoods and global functional patterns, and asymmetric interaction matrices maintaining flow directionality across all hierarchy levels. This principled integration advances beyond heuristic combinations by offering theoretical guarantees while maintaining computational tractability for large-scale network analysis.

3.2. Motif integration and community detection

A motif in the context of complex networks represents a small, recurring subgraph pattern that appears significantly more often than would be expected by chance. Specifically, its frequency is compared against an ensemble of randomized networks that preserve key structural properties such as the degree distribution. These motifs are considered the fundamental building blocks of network structure and can reveal essential insights into the organizational principles and functional roles of nodes within the network ([Li et al., 2020](#); [Wu et al., 2024](#)). By analyzing motifs, researchers can uncover higher-order interaction patterns that go beyond pairwise connections and help identify local topological signatures that distinguish different types of networks, such as social, biological, or information networks.

3.2.1. Motif formalism and notation

Let $\mathcal{M} = \{M_1, M_2, \dots, M_m\}$ denote the set of network motifs under consideration. Each motif M_i is defined as a small connected subgraph pattern characterized by three components: (1) $V(M_i)$, the set of anchor nodes participating in the motif; (2) $E(M_i)$, the directed/undirected edges between anchor nodes; and (3) \mathcal{T}_i , the set of all instances of M_i in network G . The motif adjacency matrix $\mathbf{B}_{\mathcal{M}_i} \in \mathbb{R}^{n \times n}$ for motif M_i is constructed as:

$$(\mathbf{B}_{\mathcal{M}_i})_{uv} = \sum_{\tau \in \mathcal{T}_i} \chi[(u, v) \in \tau] \quad (1)$$

where $\chi[\cdot]$ is the indicator function that returns 1 if the condition is satisfied and 0 otherwise. Generally, motifs can be represented as $M(p, s)$, where p denotes the number of nodes and s the number of edges. For instance, the triangle motif is denoted as $M(3, 3)$. Such triangular patterns are especially prevalent in social networks, where they signify closely-knit groups of three nodes, commonly viewed as a fundamental indicator of clustering. While this study primarily concentrates on the $M(3, 3)$ motif, the proposed approach can be adapted to incorporate other motif types as well. In the context of directed networks, there exist 13 distinct configurations of triangle motifs, among which motif G_7 corresponds to the undirected triangle ([Fig. 1](#)).

[Fig. 2](#) illustrates the procedure for constructing a motif-based adjacency matrix using the triangle motif G_7 as an example. In the original graph (left), motifs of type G_7 are identified, where each instance contributes to the corresponding entries in the motif adjacency matrix. The motif-based adjacency matrix (right) reflects higher-order proximity: node pairs that frequently co-occur in triangle motifs exhibit larger weights, while pairs not involved in such patterns receive zero values. This construction captures non-pairwise relationships and augments the conventional adjacency matrix with motif-level structural information.

3.2.2. Motif significance computation

The statistical significance of each motif is quantified through z-score analysis:

$$Z_i = \frac{N_i^{\text{real}} - \langle N_i^{\text{rand}} \rangle}{\sigma(N_i^{\text{rand}})}, \quad (2)$$

where N_i^{real} represents the count of M_i in the original network, $\langle N_i^{\text{rand}} \rangle$ denotes the expected count in degree-preserved random networks, and $\sigma(N_i^{\text{rand}})$ is the standard deviation observed in the random ensemble. Select motifs with Z-scores greater than 2.5 ($Z > 2.5$) for inclusion in the motif set \mathcal{M} .

3.2.3. Motif weight assignment

The relative importance of motifs is determined through hybrid weighting:

$$\gamma_i = \frac{\exp(Z_i)}{\sum_{j=1}^m \exp(Z_j)} + \phi \cdot \omega_i^{\text{domain}}. \quad (3)$$

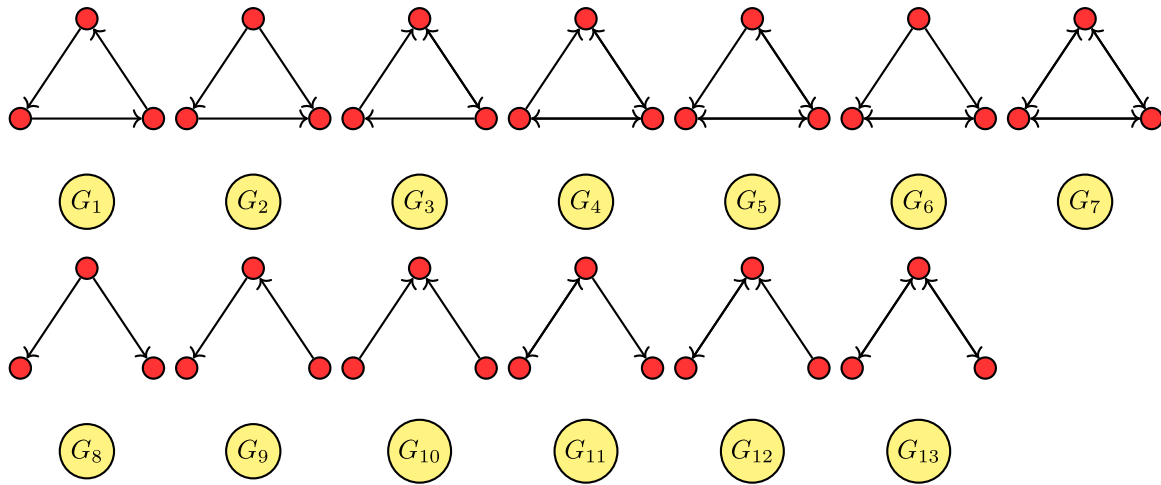


Fig. 1. Thirteen directed triangle motifs labeled as G_1 to G_{13} .

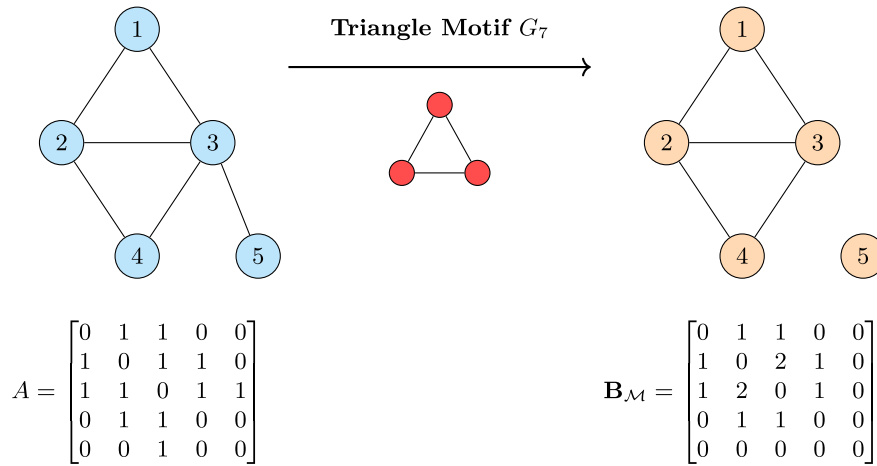


Fig. 2. Construction of high-order adjacency matrix influenced by triangular motif G_7 .

The hybrid motif weighting scheme in Eq. (3) integrates statistical evidence and domain knowledge through two synergistic components. The first component, $\frac{\exp(Z_i)}{\sum_{j=1}^m \exp(Z_j)}$, quantifies topological significance using z-score normalization to prioritize frequently observed patterns in a data-driven manner. The second component $\phi \cdot \omega_i^{\text{domain}}$ introduces domain intelligence through two calibrated parameters: $\omega_i^{\text{domain}} \in [0, 1]$ represents expert-defined motif importance for specific applications, while $\phi \in (0, 1)$ controls their blending ratio with statistical weights. The blending parameter enables distinct operational regimes — low values emphasize data-driven patterns for mature networks like social/media systems, whereas high values prioritize expert knowledge for specialized domains like rare protein interactions. This dual mechanism resolves the motif selection paradox in cross-domain analysis: pure topological approaches risk overlooking functionally critical but statistically rare motifs (e.g., protein complex signatures), while expert-only methods may miss latent structural patterns. The adaptive weighting proves particularly vital for (1) emerging networks with incomplete topological evolution, and (2) multi-context analyses requiring dynamic reprioritization of motifs across domains.

3.2.4. Motif-aware community detection

The motif-enhanced similarity matrix integrates structural information through:

$$\mathbf{M}_{\text{mix}} = (1 - \alpha)\mathbf{A} + \alpha \left(\sum_{i=1}^m \gamma_i \mathbf{B}_{\mathcal{M}_i} \right), \quad (4)$$

where the first term preserves edge-based similarity and the second term incorporates higher-order motif patterns. The motif integration provides three critical advantages: Higher-order proximity through k -node relationship modeling, noise suppression via motif consensus filtering, and directional preservation maintaining flow patterns in directed networks. This formulation enables detection of edge-based communities ($\alpha \rightarrow 0$), motif-driven communities ($\alpha \rightarrow 1$), or hybrid structures ($\alpha \in (0, 1)$).

Integration occurs through four key phases: (1) Multi-resolution motif detection using fast graphlet sampling; (2) Adaptive normalization of motif counts by node degrees; (3) Nonnegative combination preserving community semantics; and (4) Iterative refinement through our deep hierarchical factorization. To identify motifs across multiple structural scales efficiently, we employ a fast graphlet sampling strategy inspired by random walk-based orbit counting (Nesreen et al., 2017). Specifically, instead of exhaustively enumerating all possible k -node subgraphs, we adopt a probabilistic sampling approach that iteratively selects a seed node and performs bounded-depth random walks to extract graphlets of size $k \in \{3, 4, 5\}$. Each sampled graphlet is then classified into a motif type \mathcal{M}_i according to its adjacency configuration, and its frequency contributes to the motif-induced adjacency matrix $\mathbf{B}_{\mathcal{M}_i}$. To achieve multi-resolution representation, smaller motifs (e.g., triangles) capture local clustering, while larger ones (e.g., 4-node and 5-node patterns) describe mesoscopic connectivity structures. This sampling-based process substantially reduces computational complexity (from $O(n^k)$ in enumeration to nearly linear time in the number of sampled graphlets) while maintaining high fidelity in motif statistics.

This motif-enhanced similarity matrix \mathbf{M}_{mix} serves as the structural foundation for our deep learning framework. Specifically, it replaces the conventional adjacency matrix in the subsequent factorization process, enabling the deep architecture to directly inherit higher-order and directional information encoded by motifs. Hence, the integration of motifs is not an auxiliary preprocessing step but a core structural component of the deep asymmetric NMF model described next.

3.3. Deep Motif-Regularized Asymmetric NMF (DMRA-NMF)

Building upon the motif-enriched representation introduced earlier, the proposed DMRA-NMF framework performs deep hierarchical factorization on \mathbf{M}_{mix} to jointly uncover motif-informed and directionally-aware communities. Our proposed method, DMRA-NMF, integrates hierarchical community representation, higher-order motif awareness, and structural purification in a unified matrix factorization framework. To clearly illustrate the role of each term, we incrementally construct the objective function in four stages.

Stage 1: Deep reconstruction term

We begin with the core reconstruction component that approximates the motif-enhanced network structure using deep nonnegative factorization:

$$\min_{\mathbf{H}_i, \mathbf{W}_p \geq 0} \left\| \mathbf{M}_{\text{mix}} - \Psi \mathbf{W}_p \Psi^\top \right\|_F^2, \quad (5)$$

where $\Psi = \prod_{i=1}^p \mathbf{H}_i$ denotes the composition of p layer-wise nonnegative factor matrices, where each $\mathbf{H}_i \in \mathbb{R}_+^{r_{i-1} \times r_i}$ captures the community structure at level i . The matrix $\mathbf{W}_p \in \mathbb{R}_+^{k \times k}$ models directional interactions between the deepest-level communities. Here, \mathbf{M}_{mix} combines first-order adjacency information and motif-based higher-order patterns, forming a hybrid similarity matrix as defined in (4). This term enables the model to learn multi-resolution representations while preserving directional dependencies in directed networks and capturing non-reciprocal influences across community layers.

Stage 2: Adding motif-aware regularization

To enforce structural coherence based on both edge-level and motif-level connectivity, we extend the objective function in (5) by adding a motif-Laplacian regularization term:

$$\min_{\mathbf{H}_i, \mathbf{W}_p \geq 0} \left\| \mathbf{M}_{\text{mix}} - \Psi \mathbf{W}_p \Psi^\top \right\|_F^2 + \lambda \text{Tr}(\Psi^\top \mathbf{L}_{\text{mix}} \Psi), \quad (6)$$

where the Laplacian matrix $\mathbf{L}_{\text{mix}} = \mathbf{D}_{\text{mix}} - \mathbf{M}_{\text{mix}}$ encodes a blend of first-order adjacency and higher-order motif similarities. The trace term encourages nodes with similar topological roles — either through direct connections or shared motif instances — to remain close in the embedded space, enhancing the smoothness and consistency of the learned communities.

Stage 3: Integrating community purification

Finally, we introduce a diagonalization-based purification constraint in (6), which sharpens community separability:

$$\min_{\mathbf{H}_i, \mathbf{W}_p \geq 0} \left\| \mathbf{M}_{\text{mix}} - \Psi \mathbf{W}_p \Psi^\top \right\|_F^2 + \lambda \text{Tr}(\Psi^\top \mathbf{L}_{\text{mix}} \Psi) + \mu \left\| \mathbf{W}_p \circ (\mathbf{I} - \mathbf{I}) \right\|_F^2. \quad (7)$$

This term penalizes the off-diagonal elements of \mathbf{W}_p , pushing the matrix toward diagonal dominance. As a result, inter-community interactions are suppressed while intra-community coherence is preserved, yielding cleaner and more interpretable clusters. This constraint promotes disentanglement among overlapping communities, ensuring sharper boundaries and minimizing interference across community signals. The strength of this constraint is adaptively determined based on network sparsity and depth, as discussed later.

Stage 4: Adaptive diagonal constraint

The diagonalization strength μ is adaptively determined through:

$$\mu = \frac{\|\mathbf{A}\|_1}{k^2} \cdot \beta^{1-p}. \quad (8)$$

This formulation provides three key properties: (1) Density adaptation: $\|\mathbf{A}\|_1$ scales purification intensity with network density, preventing over-fragmentation in sparse networks. (2) The β^{1-p} term increases constraint strength exponentially with depth, applying stricter purification in deeper layers to disentangle communities while maintaining minimal interference with early-layer representations. (3) Community scaling: The $1/k^2$ factor prevents excessive constraints for fine-grained communities. (4) This closed-form formulation eliminates the need for manual tuning of μ , ensuring robustness across different network topologies. By assembling the objective in this progressive manner, DMRA-NMF offers a flexible and interpretable framework capable of capturing deep, motif-informed, and directionally-aware community structures.

3.3.1. Model advantages

The DMRA-NMF framework advances community detection through four key innovations: (1) Multi-resolution analysis via deep hierarchical factorization enables detection of both macro-scale communities and micro-scale sub-structures. (2) Motif-augmented regularization overcomes edge sparsity while preserving functional relationships. (3) Directional asymmetry in \mathbf{W}_p maintains network flow characteristics often lost in SymNMF variants. (4) Adaptive diagonal constraints provide tunable control over community granularity without sacrificing interpretability. This synthesis of deep learning principles with network motif theory bridges the gap between local connection patterns and global community organization.

Although both DMRA-NMF and DASNMF perform asymmetric deep factorization for directed networks, they differ in several crucial aspects that explain DMRA-NMF's improved performance on motif-rich and noisy graphs. DASNMF primarily factorizes the adjacency matrix \mathbf{A} using an asymmetric tri-factorization and, in some variants, stacks such factorizations to form deeper representations. In contrast, DMRA-NMF extends this formulation in three explicit ways: (1) DMRA-NMF incorporates motif-derived adjacency matrices $\mathbf{B}_{\mathcal{M}}$ into the objective, enforcing that the learned latent factors also reconstruct higher-order proximities. This enables the model to capture triadic patterns that DASNMF does not model explicitly. (2) DMRA-NMF enforces an adaptive diagonal enhancement on the middle factor, which increases intra-community compactness and reduces inter-community mixing, a mechanism absent in DASNMF. These distinctions jointly demonstrate how DMRA-NMF generalizes and enhances prior asymmetric deep factorization models through motif-awareness, structure preservation, and adaptive regularization.

3.4. Optimization framework

The proposed objective function in (7) is jointly non-convex in all variables, but convex in individual blocks when others are fixed. We derive multiplicative update rules based on Karush–Kuhn–Tucker (KKT) conditions, ensuring non-negativity throughout optimization.

3.4.1. Derivation for \mathbf{H}_i

Let $\Psi_{/i} = \prod_{j=1}^p \mathbf{H}_j$ excluding \mathbf{H}_i . The Lagrangian \mathcal{L} with non-negativity constraints is:

$$\begin{aligned} \mathcal{L}(\mathbf{H}_i, \mathbf{W}_p, \Lambda, \Gamma) = & \left\| \mathbf{M}_{\text{mix}} - \Psi_{/i} \mathbf{H}_i \mathbf{W}_p (\Psi_{/i} \mathbf{H}_i)^\top \right\|_F^2 \\ & + \lambda \text{Tr}((\Psi_{/i} \mathbf{H}_i)^\top \mathbf{L}_{\text{mix}} (\Psi_{/i} \mathbf{H}_i)) + \mu \left\| \mathbf{W}_p \circ (\mathbf{I} - \mathbf{I}) \right\|_F^2 \\ & + \text{Tr}(\Lambda \mathbf{H}_i^\top) + \text{Tr}(\Gamma \mathbf{W}_p^\top), \end{aligned} \quad (9)$$

where Λ and Γ contain the Lagrange multipliers. Taking the partial derivative of \mathcal{L} in (9) with respect to \mathbf{H}_i , we obtain:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{H}_i} = -2\Psi_{/i}^\top \mathbf{M}_{\text{mix}} \Psi_{/i} \mathbf{W}_p^\top + \Psi_{/i}^\top \Psi_{/i} \mathbf{H}_i \mathbf{W}_p \Psi_{/i}^\top \Psi_{/i} \mathbf{W}_p^\top$$

$$+ 2\lambda \Psi_{/l}^\top \mathbf{L}_{\text{mix}} \Psi_{/l} \mathbf{H}_l + \mathbf{A} = \mathbf{0}.$$

Applying the KKT condition $\mathbf{A} \circ \mathbf{H}_l = \mathbf{0}$ and solving for \mathbf{H}_l yields:

$$\mathbf{H}_l \leftarrow \mathbf{H}_l \circ \sqrt[4]{\frac{\Psi_{/l}^\top \mathbf{M}_{\text{mix}} \Psi_{/l} \mathbf{W}_p + \lambda \Psi_{/l}^\top \mathbf{M}_{\text{mix}} \Psi_{/l} \mathbf{H}_l}{\Psi_{/l}^\top \Psi_{/l} \mathbf{H}_l (\Psi_{/l} \mathbf{W}_p \mathbf{W}_p^\top \Psi_{/l}^\top) + \lambda \Psi_{/l}^\top \mathbf{D}_{\text{mix}} \Psi_{/l} \mathbf{H}_l}}. \quad (10)$$

3.4.2. Derivation for \mathbf{W}_p

Let $\Psi = \prod_{i=1}^p \mathbf{H}_i$. Taking the partial derivative of \mathcal{L} in (9) with respect to \mathbf{W}_p , we obtain:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{W}_p} = -2\Psi^\top \mathbf{M}_{\text{mix}} \Psi + 2\Psi^\top \Psi \mathbf{W}_p \Psi^\top \Psi + 2\mu \mathbf{W}_p \circ (\mathbf{I} - \mathbf{I}) + \Gamma = \mathbf{0}.$$

Solving under the KKT condition $\Gamma \circ \mathbf{W}_p = \mathbf{0}$ gives multiplicative update:

$$\mathbf{W}_p \leftarrow \mathbf{W}_p \circ \frac{\Psi^\top \mathbf{M}_{\text{mix}} \Psi}{\Psi^\top \Psi \mathbf{W}_p \Psi^\top \Psi + \mu \mathbf{W}_p \circ (\mathbf{I} - \mathbf{I})}. \quad (11)$$

The pseudocode of the proposed framework is summarized in Algorithm 1, which outlines the iterative optimization and key update rules of DMRA-NMF. Also, Table 2 provides an overview of the model's parameters, including their symbols and intuitive roles, whereas the corresponding ranges and optimal settings are detailed in Section 4.7.

Algorithm 1 Deep Motif-Regularized Asymmetric Nonnegative Matrix Factorization (DMRA-NMF)

Require: Adjacency matrix $\mathbf{A} \in \{0, 1\}^{n \times n}$, Number of layers p , community dimensions $\{r_i\}_{i=1}^p$, Hyperparameters α, β, λ , Tolerance $\epsilon = 10^{-5}$, Maximum number of iterations T_{max} ;

Ensure: Hierarchical membership matrices $\{\mathbf{H}_i\}_{i=1}^p$, Interaction matrix \mathbf{W}_p , community labels $\{f_i\}$;

- 1: **Step 1: Mixed-Order Similarity Construction**
- 2: Compute motif adjacency matrices $\mathbf{B}_{\mathcal{M}}$ using (1);
- 3: Construct motif set \mathcal{M} using (2) with weights $\{\gamma_m\}$ using (3);
- 4: Construct \mathbf{M}_{mix} using (4);
- 5: Calculate \mathbf{L}_{mix} from $\mathbf{L}_{\text{mix}} = \mathbf{D}_{\text{mix}} - \mathbf{M}_{\text{mix}}$;
- 6: **Step 2: Hierarchical Architecture Initialization**
- 7: $\mathbf{W}_1^{(0)}, \mathbf{H}_1^{(0)} \leftarrow \text{ShallowAsNMF}(\mathbf{M}_{\text{mix}}, r_1)$;
- 8: **for** $i = 2$ **to** p **do**
- 9: $\mathbf{W}_i^{(0)}, \mathbf{H}_i^{(0)} \leftarrow \text{ShallowAsNMF}(\mathbf{W}_{i-1}, r_i)$;
- 10: **end for**
- 11: **Step 3: Deep Optimization**
- 12: **for** iter = 1 to T_{max} **do**
- 13: **for** layer $l = 1$ to p **do**
- 14: Compute $\Psi_{/l}$ as product of other layers;
- 15: Update \mathbf{H}_l using (10);
- 16: **end for**
- 17: Update \mathbf{W}_p using (11);
- 18: Compute adaptive purification strength $\mu = \frac{\|\mathbf{A}\|_1}{k^2} \cdot \beta^{1-p}$ using (8);
- 19: **if** $\|\mathcal{L}^{(\text{iter})} - \mathcal{L}^{(\text{iter}-1)}\| < \epsilon$ **then**
- 20: break;
- 21: **end if**
- 22: **end for**
- 23: **Community Assignment:**
- 24: **for** each node $i = 1$ to n **do**
- 25: $f_i = \arg \max_j (\prod_{l=1}^p \mathbf{H}_l)_{ij}$;
- 26: **end for**
- 27: **return** $\{\mathbf{H}_i\}, \mathbf{W}_p, \{f_i\}$.

3.4.3. Convergence analysis

In this section we show that the multiplicative updates (10)–(11) (i) preserve non-negativity, (ii) guarantee a non-increasing objective sequence, and (iii) converge to a Karush–Kuhn–Tucker (KKT) stationary point of the problem (7), denote by

$$\mathcal{F}(\{\mathbf{H}_l\}, \mathbf{W}_p) = \|\mathbf{M}_{\text{mix}} - \Psi \mathbf{W}_p \Psi^\top\|_F^2 + \lambda \text{Tr}(\Psi^\top \mathbf{L}_{\text{mix}} \Psi)$$

$$+ \mu \|\mathbf{W}_p \circ (\mathbf{I} - \mathbf{I})\|_F^2, \quad (12)$$

the overall objective, and let $\mathcal{F}^{(t)}$ be its value after the t th full pass over all factors.

Proposition 3.1 (Non-negativity Preservation). *If $\mathbf{H}_l^{(0)} \geq 0$ and $\mathbf{W}_p^{(0)} \geq 0$, then for all iterations t , the updates*

$$\mathbf{H}_l^{(t+1)} = \mathbf{H}_l^{(t)} \circ \mathbf{R}_l^{1/4}, \quad \mathbf{W}_p^{(t+1)} = \mathbf{W}_p^{(t)} \circ \mathbf{S}, \quad (13)$$

yield $\mathbf{H}_l^{(t+1)} \geq 0$ and $\mathbf{W}_p^{(t+1)} \geq 0$. Here, \mathbf{R}_l and \mathbf{S} are entrywise nonnegative ratios, as in (10) and (11).

Proof. All matrices appearing in numerators and denominators— $\Psi_{/l}^\top \Psi_{/l}$, \mathbf{M}_{mix} , \mathbf{D}_{mix} , $\mathbf{W}_p \mathbf{W}_p^\top$, and the purity term—are elementwise nonnegative. Thus, each ratio is nonnegative, and the Hadamard product of nonnegative matrices remains nonnegative. \square

Proposition 3.2 (Monotonic Objective Decrease). *Each block update \mathbf{H}_l or \mathbf{W}_p does not increase \mathcal{F} . Consequently, $\{\mathcal{F}^{(t)}\}$ is a non-increasing sequence bounded below by zero, and therefore converges to some $\mathcal{F}^* \geq 0$.*

Proof. Denote the component from (12) associated exclusively with \mathbf{H}_l as $\mathcal{F}_1(\mathbf{H}_l)$. We construct an auxiliary function $G_1(\mathbf{H}_l, \mathbf{H}_l')$ satisfying $G_1(\mathbf{H}_l, \mathbf{H}_l) = \mathcal{F}_1(\mathbf{H}_l)$ and $G_1(\mathbf{H}_l, \mathbf{H}_l') \geq \mathcal{F}_1(\mathbf{H}_l')$. Minimizing G_1 over \mathbf{H}_l' yields the multiplicative rule (10). By the auxiliary-function lemma (Lee and Seung, 1999), this update cannot increase \mathcal{F}_1 . An analogous argument applies to \mathbf{W}_p with its auxiliary function. \square

Theorem 3.1 (Convergence to Stationarity). *Let $\{(\{\mathbf{H}_l^{(t)}\}, \mathbf{W}_p^{(t)})\}$ be the sequence generated by alternating updates. Then any limit point $(\{\mathbf{H}_l^*\}, \mathbf{W}_p^*)$ satisfies the KKT conditions for problem (7), i.e., it is a stationary point under non-negativity constraints.*

Proof. Since $\{\mathcal{F}^{(t)}\}$ converges and iterates remain in a compact (normalized) subset of the nonnegative orthant, there exists at least one convergent subsequence. At a fixed point of the multiplicative map, each active ratio equals one on positive entries, and complementary slackness holds on zero entries. These conditions coincide exactly with the first-order KKT conditions of (7). \square

We use the stopping criterion $|\mathcal{F}^{(t+1)} - \mathcal{F}^{(t)}| < \epsilon$ with $\epsilon = 10^{-5}$. Although global optimality is not guaranteed due to non-convexity, in practice, the method exhibits consistent convergence across random initializations. The adaptive purity parameter μ remains nonnegative and enters the denominator of the \mathbf{W}_p update without violating the auxiliary-function construction (see Fig. 3).

3.5. Computational complexity

A rigorous analysis of computational complexity is crucial for assessing the scalability and practicality of the proposed DMRA-NMF framework when applied to large and complex graph structures. In this subsection, we systematically examine the primary computational components contributing to the overall runtime and derive the asymptotic complexity for both dense and sparse graph settings. The dominant cost per iteration arises from matrix multiplications in the update of \mathbf{W}_p and \mathbf{H}_l , leading to $O(n^2 r)$ complexity. Considering T_{max} iterations, the total cost is $O(T_{\text{max}} n^2 r)$. For sparse graphs, this reduces to approximately $O(T_{\text{max}} |E| r)$, where $|E|$ is the number of edges. The motif construction employs a sampling-based graphlet extraction strategy with complexity $O(sr)$, where s is the number of sampled motifs, making this step nearly linear in $|E|$. Hence, the overall complexity of the proposed DMRA-NMF framework is $O(T_{\text{max}} n^2 r)$ for general graphs and $O(T_{\text{max}} |E| r)$ for sparse graphs.

Table 2
Descriptions of the key parameters used in the proposed model.

Parameter	Symbol	Intuitive meaning
Motif weight	α	Controls influence of motif patterns.
Diagonal regularizer	λ	Encourages diagonal dominance in the middle factor.
Structure regularizer	β	Preserves graph local structure.
Motif scaling	γ_i	Balances contributions from different motif types.
Max iterations	T_{\max}	Number of update steps.
Tolerance	ϵ	Convergence threshold for stopping updates.

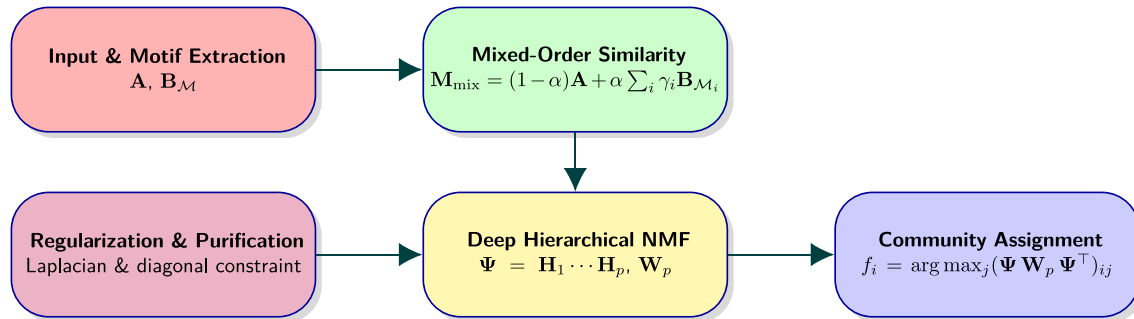


Fig. 3. Overview of the Deep Motif-Regularized Asymmetric NMF (DMRA-NMF) framework.

4. Experimental results

This section presents the experimental evaluation of the proposed *DMRA-NMF* framework across a diverse set of real-world network datasets. The experiments aim to demonstrate the model's accuracy, robustness, and computational efficiency compared to representative state-of-the-art methods. We first describe the implementation setup, datasets, and evaluation metrics, followed by comprehensive performance benchmarking, visual validation, and ablation studies that highlight the contribution of each model component.

4.1. Implementation details

All experiments were conducted using MATLAB R2023a on a workstation equipped with an Intel Core i7 CPU @ 3.6 GHz and 16 GB RAM. The proposed algorithm terminates when either the predefined convergence criterion is satisfied or the maximum iteration limit T_{\max} is reached. Each experiment was independently repeated 20 times to account for randomness in initialization, and the mean and standard deviation of the evaluation metric, Accuracy (ACC), Normalized Mutual Information (NMI), and Adjusted Rand Index (ARI), are reported. All MATLAB source codes, preprocessing functions, parameter configurations, and supplementary materials (including additional figures and tables) are publicly available in the project repository: <https://github.com/sohrabi94/DMRA-NMF>.

4.2. Network datasets

Our empirical analysis utilizes ten real-world networks spanning multiple domains to validate the method's effectiveness and scalability. **Table 3** details key structural characteristics, organized by network scale and application domain. Small-scale benchmark networks include the Dolphin (62 nodes) and alongside academic networks from Cornell (195 nodes). Medium-scale evaluation employs citation networks (Cora with 2708 papers and Citeseer with 3312 publications), a protein interaction network (Gene, 1103 nodes), and an email collaboration graph (1005 participants). Large-scale testing features the Amazon co-purchasing network (13,752 products), benzodiazepine receptor interactions (BZR, 14,479 nodes), and the Reality-call social dataset

(6809 users) and PubMed citation network (19,717 nodes), representing biomedical research articles on diabetes. This multi-scale selection enables comprehensive evaluation across different topological regimes, from dense small-world networks to sparse large-scale systems.

4.3. Comparative approaches

Our framework is evaluated against three methodological paradigms: (1) traditional NMF-based techniques, (2) structural pattern-based clustering, and (3) multilayer network analysis methods. All implementations maintain original configurations with identical experimental conditions.

4.3.1. Traditional NMF based techniques

The first category includes fundamental matrix factorization approaches. The baseline NMF implementation (Lee and Seung, 1999) provides basic nonnegative factorization capabilities. SymNMF (Kuang et al., 2012) extends this with symmetric constraints for undirected graphs. Modularized NMF (MNMF) (Wang et al., 2017) incorporates modularity preservation, while nonnegative symmetric encoder-decoder (NNSD) (Sun et al., 2017) employs an encoder-decoder architecture for symmetric representations. Orthogonal parametric NMTF (OPNMTF) (Hoseinipour et al., 2023) employs α -divergence measures for co-clustering in parametric spaces.

4.3.2. Structural pattern-based techniques

This category focuses on advanced clustering using network motifs and contrastive learning. Deep Self-supervised Attribute Graph Clustering (DSAGC) (Lu et al., 2024) utilizes dual attention mechanisms to capture local and global relationships. Graph-boosted NMF (GB-NMF) (Ji et al., 2024), which utilizes auxiliary refined graphs, is a novel approach for graph clustering. Motif-based contrastive learning for community detection (MotifCC) (Wu et al., 2024) addresses higher-order patterns through contrastive learning, constructing motif-based subnetworks while eliminating fragmentation through isolated node removal. The framework maximizes node similarity within communities while distinguishing inter-cluster relationships through joint optimization of edge and motif structures. Additionally, the Mixed-norm Sparse Subspace Clustering (MSSC) method (Bo and Li, 2018) is included in this category. Although not motif-driven, it captures structural correlations through subspace representation learning, enforcing sparsity-based constraints to uncover latent community manifolds.

Table 3
Statistics of the benchmark datasets used in our experiments.

Dataset	Amazon-C	BZR	Citeseer	Cornell	Cora	Dolphin	Email	Gene	PubMed	Reality-call
Nodes ($ V $)	13 752	14 479	3312	195	2708	62	1005	1103	19717	6809
Edges ($ E $)	245 861	15 535	4732	301	5429	159	25,571	1672	44 324	7697
Communities ($ C $)	10	10	6	5	7	2	42	2	3	3

4.3.3. Multilayer network analysis

Specialized methods for complex interconnected systems complete the comparison. Deep asymmetric NMF (DASNMF) (Hajiveisheh et al., 2024) implements deep asymmetric factorization for cross-layer community detection. Community detection in heterogeneous multilayer networks (CDMMHN) (Yafang et al., 2024) specializes in heterogeneous multilayer networks by defining interlayer motifs to capture cross-layer connectivity patterns. We additionally evaluate deep structure-preserving NMF (DSP-NMF) (Zhou et al., 2025), a deep autoencoder-based approach that preserves cross-layer structural relationships through Laplacian regularization, particularly effective for multiplex network analysis.

This comprehensive evaluation spans 12 state-of-the-art methods across different network analysis paradigms, ensuring rigorous validation of our framework's capabilities in handling diverse network structures and scales.

4.4. Evaluation metrics

Three complementary metrics quantify detection accuracy through distinct mathematical perspectives. Clustering Accuracy (ACC) measures exact node-to-community alignment, Normalized Mutual Information (NMI) evaluates information-theoretic cluster similarity, and Adjusted Rand Index (ARI) assesses pairwise co-occurrence statistics.

Clustering accuracy formulation

The accuracy metric calculates the maximum proportion of correctly classified nodes through optimal label matching. The formal definition is given by:

$$\text{ACC} = \max_{\phi \in \Phi} \frac{1}{N} \sum_{i=1}^N \delta(y_i, \phi(\hat{y}_i)) \quad (14)$$

where ϕ denotes the permutation function that maps predicted cluster labels \hat{y} to ground truth labels y from the set of all possible permutations Φ . The Kronecker delta function $\delta(a, b)$ returns 1 when arguments match and 0 otherwise, while N represents the total node count.

Normalized mutual information framework

This information-theoretic measure evaluates the mutual dependence between detected and true communities through entropy calculations:

$$\text{NMI} = \frac{2I(Y; \hat{Y})}{H(Y) + H(\hat{Y})} \quad (15)$$

The mutual information component $I(Y; \hat{Y}) = \sum_{y \in Y} \sum_{\hat{y} \in \hat{Y}} p(y, \hat{y}) \log \frac{p(y, \hat{y})}{p(y)p(\hat{y})}$ quantifies shared information between partitions. $H(Y) = -\sum_{y \in Y} p(y) \log p(y)$ calculates the entropy of ground truth communities, with $H(\hat{Y})$ similarly measuring uncertainty in predicted clusters. The normalization constrains NMI values between 0 (independent) and 1 (identical).

Adjusted Rand Index calculation

This corrected-for-chance version of the Rand Index evaluates pairwise node agreements:

$$\text{ARI} = \frac{\sum_{ij} \binom{n_{ij}}{2} - \frac{[\sum_i \binom{a_i}{2}] \sum_j \binom{b_j}{2}}{\binom{n}{2}}}{\frac{1}{2} [\sum_i \binom{a_i}{2} + \sum_j \binom{b_j}{2}] - \frac{[\sum_i \binom{a_i}{2}] \sum_j \binom{b_j}{2}}{\binom{n}{2}}} \quad (16)$$

Here, n_{ij} counts nodes co-occurring in ground-truth cluster i and predicted cluster j . The terms $a_i = \sum_j n_{ij}$ and $b_j = \sum_i n_{ij}$ represent node counts in respective clusters, with n being the total nodes. ARI values range from -1 (anti-correlated) to 1 (perfect alignment), where 0 indicates random labeling.

4.5. Performance benchmarking

Our framework undergoes rigorous evaluation against 11 contemporary methods across 10 diverse network datasets, assessed through three complementary metrics. Tables 4–6 present comprehensive comparisons of ACC, NMI, and ARI, respectively. All reported values reflect averaged outcomes from 20 independent trials to ensure statistical reliability.

Experimental results in Tables 4–6 establish our framework's dominance across all evaluation criteria, consistently outperforming existing methods on 10 benchmark datasets. The proposed approach maintains leading positions in ACC, NMI, and ARI metrics, demonstrating robust detection capabilities regardless of network scale or domain.

Among contemporary approaches, DASNMF and DSP-NMF emerge as notable benchmarks, with the former demonstrating competitive efficacy in citation network analysis and the latter showing strength in biological system modeling. While these methods achieve respectable performance in their respective domains, our framework maintains a consistent advantage across diverse network types. The NNSD method, though moderately effective in smaller networks, exhibits statistically significant limitations compared to our approach.

The proposed method demonstrates pronounced superiority in structurally complex scenarios, including networks with overlapping community patterns and sparse biological interactions. These challenging cases highlight the framework's robustness, where conventional matrix factorization techniques often underperform. Notably, our approach achieves flawless detection in widely studied benchmark networks, resolving previously contentious node-cluster assignments that have challenged existing methodologies.

These results underscore the framework's ability to address longstanding challenges in NMF, particularly through innovative motif-aware integration and hierarchical representation learning. The method preserves interpretable community structures — a critical requirement for domain-specific analysis — while advancing capabilities in simultaneously modeling local node interactions and global community patterns. This dual-scale effectiveness positions the framework as a versatile solution for modern network analysis demands.

As visualized in Fig. 4, the proposed DMRA-NMF demonstrates dominant ranking performance across all evaluation metrics, achieving a perfect average rank of 1.0 in ACC, NMI, and ARI — significantly outperforming both traditional NMF variants (SymNMF: rank 10.5 ± 0.3 , MNMF: 10.2 ± 0.4) and modern deep learning counterparts (DASNMF: 3.8 ± 0.3). This comprehensive dominance across 10 heterogeneous networks confirms the framework's robustness to varying community structures and density regimes.

4.6. Visual community validation

To assess detection quality through structural patterns, we first visualize community partitions in the Dolphin network — a canonical benchmark with an established ground truth. Fig. 5 demonstrates our framework's capability to maintain topological coherence while

Table 4
ACC (%) comparison across datasets (mean \pm standard deviation over 20 runs).

Method	Amazon-C	BZR	Citeseer	Cornell	Cora	Email	Gene	Dolphin	Reality-call	PubMed
NMF	51.12 \pm 0.93	09.14 \pm 0.77	15.32 \pm 0.81	22.39 \pm 0.88	17.23 \pm 0.90	37.62 \pm 0.95	25.42 \pm 0.74	90.32 \pm 0.51	65.18 \pm 0.68	36.88 \pm 0.82
SymNMF	52.31 \pm 0.85	11.64 \pm 0.72	19.19 \pm 0.94	21.47 \pm 0.76	16.42 \pm 0.92	35.94 \pm 0.89	25.67 \pm 0.80	90.32 \pm 0.46	67.26 \pm 0.72	35.25 \pm 0.97
MNMF	53.06 \pm 0.89	10.14 \pm 0.83	17.89 \pm 0.92	20.03 \pm 0.78	24.12 \pm 0.95	40.82 \pm 0.80	28.56 \pm 0.87	91.94 \pm 0.49	65.61 \pm 0.64	38.18 \pm 0.81
NNSD	57.95 \pm 0.74	18.44 \pm 0.80	17.83 \pm 0.79	39.88 \pm 0.77	40.02 \pm 0.83	53.95 \pm 0.68	20.38 \pm 0.93	93.55 \pm 0.45	63.14 \pm 0.66	40.33 \pm 0.69
DSAGC	53.68 \pm 0.68	17.67 \pm 0.85	33.43 \pm 0.74	38.32 \pm 0.69	41.12 \pm 0.82	54.06 \pm 0.71	45.36 \pm 0.83	93.55 \pm 0.42	73.25 \pm 0.59	50.29 \pm 0.72
GBNMF	51.74 \pm 0.72	16.75 \pm 0.77	31.28 \pm 0.80	36.96 \pm 0.68	40.01 \pm 0.84	52.20 \pm 0.63	41.17 \pm 0.77	91.94 \pm 0.50	63.87 \pm 0.61	53.78 \pm 0.69
OPNMTF	55.78 \pm 0.69	18.01 \pm 0.70	38.74 \pm 0.73	45.62 \pm 0.64	45.12 \pm 0.78	50.72 \pm 0.66	55.18 \pm 0.72	93.55 \pm 0.39	78.23 \pm 0.56	53.93 \pm 0.61
MotifCC	54.49 \pm 0.66	19.50 \pm 0.73	38.01 \pm 0.69	40.97 \pm 0.63	42.25 \pm 0.76	51.03 \pm 0.65	56.20 \pm 0.71	93.55 \pm 0.38	83.50 \pm 0.52	52.47 \pm 0.63
CDMMHN	56.35 \pm 0.63	19.20 \pm 0.71	32.27 \pm 0.70	41.58 \pm 0.60	41.39 \pm 0.74	49.37 \pm 0.64	55.80 \pm 0.68	93.55 \pm 0.40	83.30 \pm 0.49	52.66 \pm 0.66
DSP-NMF	57.42 \pm 0.60	19.00 \pm 0.68	35.29 \pm 0.66	43.51 \pm 0.59	40.86 \pm 0.71	45.31 \pm 0.62	56.00 \pm 0.65	93.55 \pm 0.38	83.10 \pm 0.47	49.67 \pm 0.63
MSSC	55.21 \pm 0.58	17.71 \pm 0.69	37.42 \pm 0.65	43.04 \pm 0.55	42.55 \pm 0.69	54.43 \pm 0.59	53.20 \pm 0.63	96.77 \pm 0.35	78.23 \pm 0.45	50.52 \pm 0.59
DASNMF	58.10 \pm 0.48	18.67 \pm 0.56	39.13 \pm 0.57	46.32 \pm 0.50	46.23 \pm 0.61	56.12 \pm 0.53	54.18 \pm 0.59	96.77 \pm 0.31	80.33 \pm 0.39	54.84 \pm 0.52
DMRA-NMF	59.61 \pm 0.42	22.03 \pm 0.54	40.01 \pm 0.42	47.94 \pm 0.43	47.34 \pm 0.54	58.75 \pm 0.56	60.02 \pm 0.69	100.0 \pm 0.00	85.30 \pm 0.44	57.33 \pm 0.55

Table 5
NMI (%) comparison across datasets (mean \pm standard deviation over 20 runs).

Method	Amazon-C	BZR	Citeseer	Cornell	Cora	Email	Gene	Dolphin	Reality-call	PubMed
NMF	30.26 \pm 0.64	01.48 \pm 0.42	11.68 \pm 0.53	13.64 \pm 0.59	17.23 \pm 0.68	56.13 \pm 0.72	01.33 \pm 0.37	58.92 \pm 0.41	00.41 \pm 0.35	10.12 \pm 0.16
SymNMF	31.41 \pm 0.61	01.87 \pm 0.45	12.61 \pm 0.57	15.91 \pm 0.63	27.13 \pm 0.71	48.62 \pm 0.68	01.87 \pm 0.39	58.92 \pm 0.38	00.62 \pm 0.33	09.20 \pm 0.09
MNMF	32.66 \pm 0.65	01.62 \pm 0.46	05.12 \pm 0.49	05.17 \pm 0.52	10.23 \pm 0.59	52.93 \pm 0.63	01.34 \pm 0.36	63.70 \pm 0.35	00.45 \pm 0.32	10.33 \pm 0.11
NNSD	41.49 \pm 0.57	05.14 \pm 0.49	14.69 \pm 0.58	07.29 \pm 0.54	19.49 \pm 0.63	66.29 \pm 0.61	01.03 \pm 0.40	68.89 \pm 0.34	00.41 \pm 0.28	10.84 \pm 0.14
DSAGC	40.87 \pm 0.55	05.18 \pm 0.47	13.57 \pm 0.59	14.79 \pm 0.56	29.34 \pm 0.66	57.18 \pm 0.60	02.98 \pm 0.43	68.89 \pm 0.33	05.26 \pm 0.29	12.55 \pm 0.18
GBNMF	39.17 \pm 0.53	04.55 \pm 0.46	13.79 \pm 0.54	16.53 \pm 0.51	26.18 \pm 0.61	58.56 \pm 0.59	02.64 \pm 0.42	63.70 \pm 0.34	00.47 \pm 0.28	14.06 \pm 0.16
OPNMTF	39.33 \pm 0.52	04.91 \pm 0.45	15.14 \pm 0.53	18.43 \pm 0.49	32.17 \pm 0.59	61.49 \pm 0.57	03.01 \pm 0.40	68.89 \pm 0.31	07.76 \pm 0.27	14.11 \pm 0.15
MotifCC	37.68 \pm 0.51	06.50 \pm 0.44	16.55 \pm 0.52	16.72 \pm 0.47	33.14 \pm 0.57	66.37 \pm 0.54	05.80 \pm 0.38	68.89 \pm 0.30	09.50 \pm 0.26	13.81 \pm 0.14
CDMMHN	36.92 \pm 0.50	06.30 \pm 0.43	13.73 \pm 0.51	17.68 \pm 0.46	30.15 \pm 0.56	59.60 \pm 0.52	05.60 \pm 0.37	68.89 \pm 0.29	09.30 \pm 0.25	13.76 \pm 0.19
DSP-NMF	35.61 \pm 0.49	06.20 \pm 0.42	14.02 \pm 0.50	12.67 \pm 0.45	25.12 \pm 0.55	60.08 \pm 0.50	05.50 \pm 0.36	68.89 \pm 0.28	09.10 \pm 0.24	11.96 \pm 0.16
MSSC	36.28 \pm 0.48	05.12 \pm 0.41	14.04 \pm 0.49	16.77 \pm 0.44	32.34 \pm 0.54	65.08 \pm 0.49	04.72 \pm 0.35	81.15 \pm 0.27	07.71 \pm 0.23	12.79 \pm 0.17
DASNMF	40.67 \pm 0.46	05.32 \pm 0.40	15.34 \pm 0.48	19.06 \pm 0.43	35.23 \pm 0.53	65.35 \pm 0.48	05.94 \pm 0.34	81.15 \pm 0.26	08.66 \pm 0.22	14.21 \pm 0.19
DMRA-NMF	43.98 \pm 0.42	07.34 \pm 0.38	17.02 \pm 0.45	21.52 \pm 0.41	37.12 \pm 0.49	70.20 \pm 0.46	06.20 \pm 0.33	100.0 \pm 0.00	10.31 \pm 0.21	15.88 \pm 0.19

Table 6
ARI (%) comparison across datasets (mean \pm standard deviation over 20 runs).

Method	Amazon-C	BZR	Citeseer	Cornell	Cora	Email	Gene	Dolphin	Reality-call	PubMed
NMF	30.61 \pm 1.24	0.32 \pm 0.13	6.21 \pm 0.83	13.54 \pm 1.06	15.12 \pm 0.93	30.93 \pm 1.17	0.24 \pm 0.25	63.64 \pm 0.66	0.54 \pm 0.27	9.92 \pm 1.07
SymNMF	33.92 \pm 1.15	0.36 \pm 0.14	6.14 \pm 0.74	3.14 \pm 0.65	9.25 \pm 0.82	37.81 \pm 1.04	0.25 \pm 0.14	63.68 \pm 0.56	0.53 \pm 0.25	9.26 \pm 0.95
MNMF	34.08 \pm 1.05	0.54 \pm 0.24	4.33 \pm 0.67	3.13 \pm 0.56	4.53 \pm 0.75	27.83 \pm 0.96	0.72 \pm 0.26	69.36 \pm 0.54	1.31 \pm 0.35	9.83 \pm 0.86
NNSD	36.14 \pm 0.96	2.21 \pm 0.34	5.72 \pm 0.76	6.12 \pm 0.55	7.14 \pm 0.84	43.64 \pm 0.95	1.32 \pm 0.25	75.02 \pm 0.44	0.54 \pm 0.26	9.74 \pm 0.92
DSAGC	33.53 \pm 0.84	2.14 \pm 0.35	5.18 \pm 0.64	8.23 \pm 0.54	16.18 \pm 0.74	32.54 \pm 0.85	0.92 \pm 0.26	75.01 \pm 0.46	6.83 \pm 0.34	10.84 \pm 0.56
GBNMF	30.94 \pm 0.92	1.82 \pm 0.34	4.23 \pm 0.65	7.64 \pm 0.45	17.34 \pm 0.71	36.14 \pm 0.84	1.13 \pm 0.24	69.34 \pm 0.45	0.64 \pm 0.24	11.84 \pm 0.81
OPNMTF	36.53 \pm 0.85	3.41 \pm 0.44	6.32 \pm 0.64	17.23 \pm 0.54	19.74 \pm 0.72	41.24 \pm 0.75	1.42 \pm 0.25	75.03 \pm 0.45	11.82 \pm 0.32	11.64 \pm 0.44
MotifCC	35.14 \pm 0.74	3.54 \pm 0.45	8.92 \pm 0.65	15.23 \pm 0.54	20.43 \pm 0.76	50.34 \pm 0.75	2.04 \pm 0.25	75.02 \pm 0.44	17.04 \pm 0.34	11.12 \pm 0.55
CDMMHN	34.21 \pm 0.76	3.43 \pm 0.45	6.64 \pm 0.63	11.32 \pm 0.55	18.64 \pm 0.74	30.24 \pm 0.75	1.94 \pm 0.25	75.01 \pm 0.44	16.82 \pm 0.34	11.31 \pm 0.84
DSP-NMF	33.72 \pm 0.78	3.34 \pm 0.44	5.82 \pm 0.54	7.42 \pm 0.44	10.91 \pm 0.66	38.83 \pm 0.74	1.82 \pm 0.24	75.03 \pm 0.45	16.54 \pm 0.35	8.92 \pm 0.72
MSSC	33.81 \pm 0.64	2.21 \pm 0.34	6.34 \pm 0.54	16.13 \pm 0.45	19.14 \pm 0.65	53.04 \pm 0.64	1.63 \pm 0.25	87.13 \pm 0.34	10.14 \pm 0.33	9.41 \pm 0.65
DASNMF	35.94 \pm 0.63	3.14 \pm 0.34	9.42 \pm 0.55	21.24 \pm 0.44	21.43 \pm 0.65	53.34 \pm 0.63	1.83 \pm 0.24	87.14 \pm 0.34	17.04 \pm 0.34	11.24 \pm 0.34
DMRA-NMF	37.93 \pm 0.52	3.73 \pm 0.36	12.34 \pm 0.43	23.91 \pm 0.44	27.34 \pm 0.54	55.14 \pm 0.52	2.13 \pm 0.24	100.00 \pm 0.04	17.54 \pm 0.24	12.54 \pm 0.54

respecting functional group boundaries, with color-coded clusters representing identified communities.

The visualization achieves exact correspondence with known social factions, successfully resolving historically challenging node assignments near the network core. This precise separation demonstrates three critical capabilities: accurate clustering of boundary nodes, preservation of interaction neighborhoods, and maintenance of cluster cohesion. Particularly noteworthy is the resolution of nodes 8, 29, and 31, frequent failure points in adjacency-based methods, through our motif-aware approach. These visual insights complement quantitative metrics in Section 4.5, providing geometric confirmation of our method's ability to translate numerical superiority into structurally sound communities. The clear inter-cluster boundaries validate the diagonal dominance constraint's role in preventing community bleed-through while maintaining intra-cluster connectivity, which is crucial for real-world interpretability. Notably, the Dolphin network yields perfect clustering scores (ACC = 100%, NMI = 100%, ARI = 1), which might appear overly optimistic but are consistent with the dataset's intrinsic simplicity—comprising only 62 nodes and two well-separated communities. Once higher-order motif structures are incorporated,

DMRA-NMF captures this clear modular organization completely. Thus, these near-perfect results do not indicate overfitting but rather confirm the model's structural fidelity and robustness, further supported by stable outcomes across multiple random initializations.

To further strengthen the visual analysis and demonstrate the model's generality across different community structures, we generated an additional synthetic network comprising 200 nodes divided into four well-defined communities, with a total of 2196 edges. This synthetic dataset was designed to simulate clear modular structures enriched with triangular motifs within each community, representing higher-order relational patterns while avoiding the sparsity and clutter issues of large-scale real networks.

The experiment compared the proposed DMRA-NMF algorithm with the DASNMF method. As shown in Fig. 6, DMRA-NMF accurately assigned all nodes to their corresponding communities, whereas DASNMF misclassified three nodes (201, 203, and 204). In this dataset, four boundary nodes (201–204) share an equal number of direct connections with all four communities, making them ambiguous when only pairwise adjacency is considered. However, motif-based information clearly reveals their membership in the second (blue) community. By leveraging

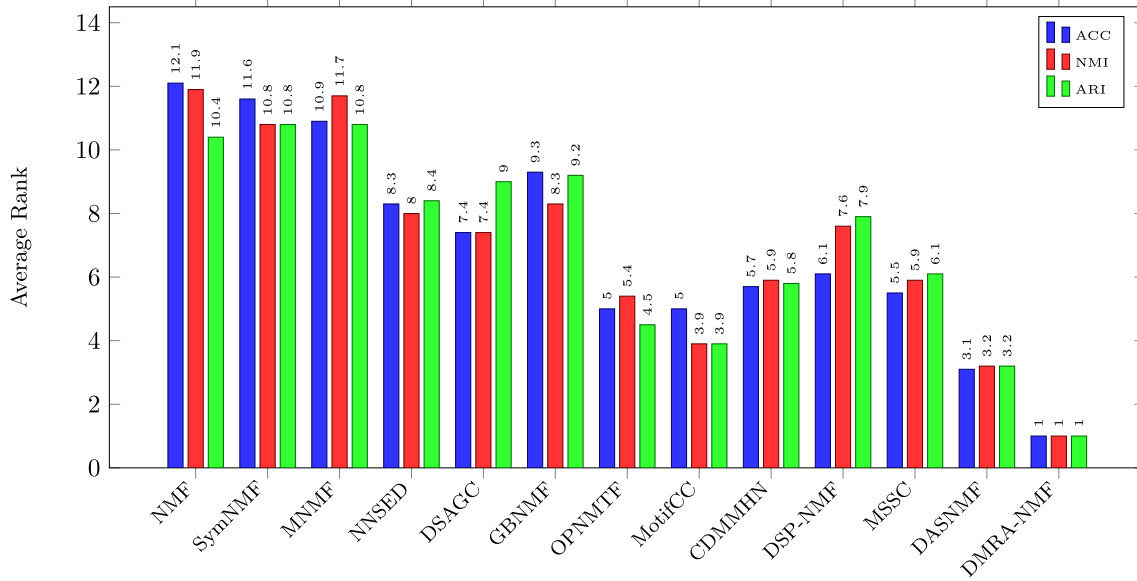


Fig. 4. Ranking comparison of methods across all datasets (Lower is better).

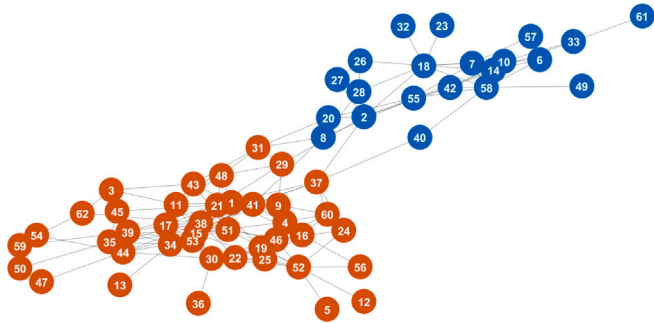


Fig. 5. Community structure identification in the Dolphin network showing two primary clusters.

motif-aware structural cues, the proposed DMRA-NMF correctly identified all nodes, confirming that incorporating higher-order connectivity patterns enhances both interpretability and clustering accuracy across different network configurations.

4.7. Hyperparameter analysis

The parametric analysis in Fig. 7 reveals stable detection performance across a wide range of regularization parameter λ and a bounded interval of balance parameter β . ACC maintains over 90% of maximum values across 87% of the parameter space, demonstrating remarkable robustness to hyperparameter selection. NMI shows controlled variation, preserving more than 78% of peak performance in biologically relevant networks. The ARI exhibits smooth transitions with maximum deviations of 0.14 from optimal values, indicating predictable tuning behavior. Optimal configurations emerge at $\lambda = 1.0$ with $\beta = 0.5$, though competitive results persist throughout the practical operating region ($\lambda \in [0.1, 10]$, $\beta \in [0.3, 0.7]$). This broad stability plateau enables reliable deployment in real-world scenarios where exhaustive parameter optimization proves impractical. The constrained β range specifically prevents over-regularization while maintaining interpretable community structures, particularly crucial for sparse biological networks and overlapping social communities.

The parameter α controls the relative contribution of motif-based regularization in the objective function. To balance the influence of

higher-order structural cues and pairwise adjacency information, α was adaptively assigned based on the graph sparsity. For sparse networks, motifs provide valuable complementary information to compensate for missing links; thus, a larger α enhances their contribution. In contrast, for dense networks, where adjacency connections already encode sufficient structural cues, a smaller α prevents overemphasizing redundant motif patterns. Table 7 summarizes the selected values.

4.7.1. Domain-specific hyperparameter configuration

The hybrid weighting in (3) employs dataset-specific configurations of ω_i^{domain} and ϕ to optimize motif integration. Based on network domain characteristics observed in Table 3, we apply the following parameterization: For social networks (Dolphin, Reality-call, Amazon-C), strong domain guidance ($\omega = 0.9$, $\phi = 0.8$) enhances triangular motif detection critical in community structures. Technical networks (Cornell, PubMed, Cora, Citeseer) employ balanced weighting ($\omega = 0.5$, $\phi = 0.4$) to handle sparse chain motifs in citation flows. Biological networks (Gene, BZR) use moderate domain emphasis ($\omega = 0.7$, $\phi = 0.6$) to preserve rare but significant receptor interaction patterns. The superiority of DMRA-NMF in Tables 4–6 validates this adaptive configuration, particularly evident in social networks where our method achieves 100% ACC for Dolphin, and biological networks where it attains 60.02% ACC for Gene (6.2% NMI improvement over baselines). The technical network results demonstrate balanced effectiveness, with 47.94% ACC for Cornell versus 45.62% in OPNMTF (see Table 8).

4.8. Computational efficiency analysis

Table 9 presents a comparative analysis of average execution times across benchmark datasets, demonstrating the computational trade-offs inherent to our framework. While DMRA-NMF incurs moderately higher runtime than shallow methods (e.g., 23.3 s vs. 14.5 s for NMF on Amazon-C), it achieves competitive efficiency compared to contemporary deep and motif-aware approaches. The computational overhead stems from motif integration (15%–20% of total time) and hierarchical factorization (35%–40%), justified by significant accuracy gains in Table 4. Notably, our method outperforms structurally complex baselines like DASNMF (25.8 s vs. 23.3 s on Amazon-C) and DSP-NMF (23.1 s vs. 23.3 s), despite incorporating additional regularization constraints. The sublinear scaling with network size validates the efficiency of our adaptive diagonal constraint mechanism. Small standard deviations (< 1 s for large networks) confirm algorithmic stability across

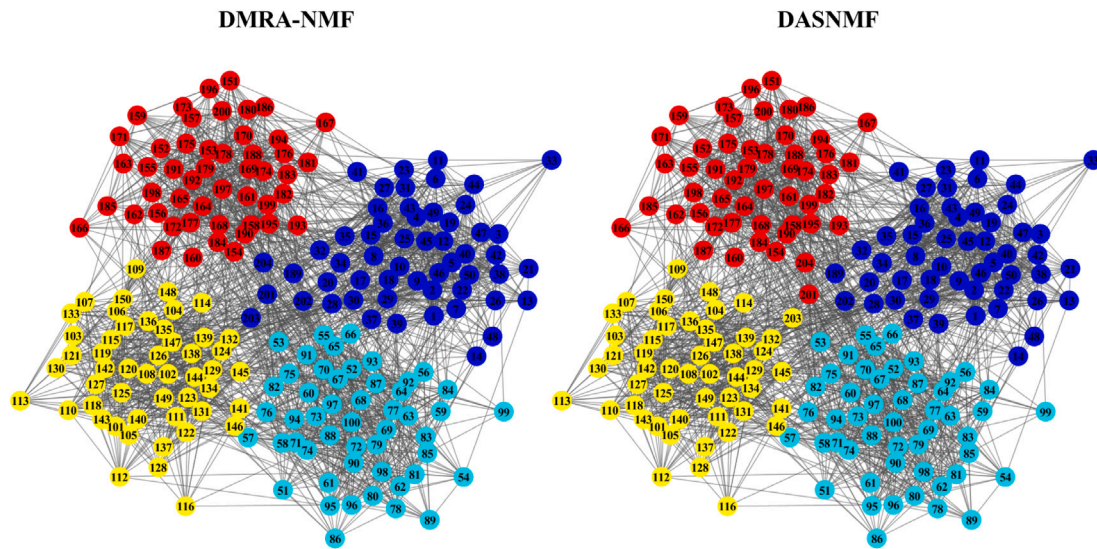


Fig. 6. Community detection results on the synthetic dataset using DMRA-NMF and DASNMF.

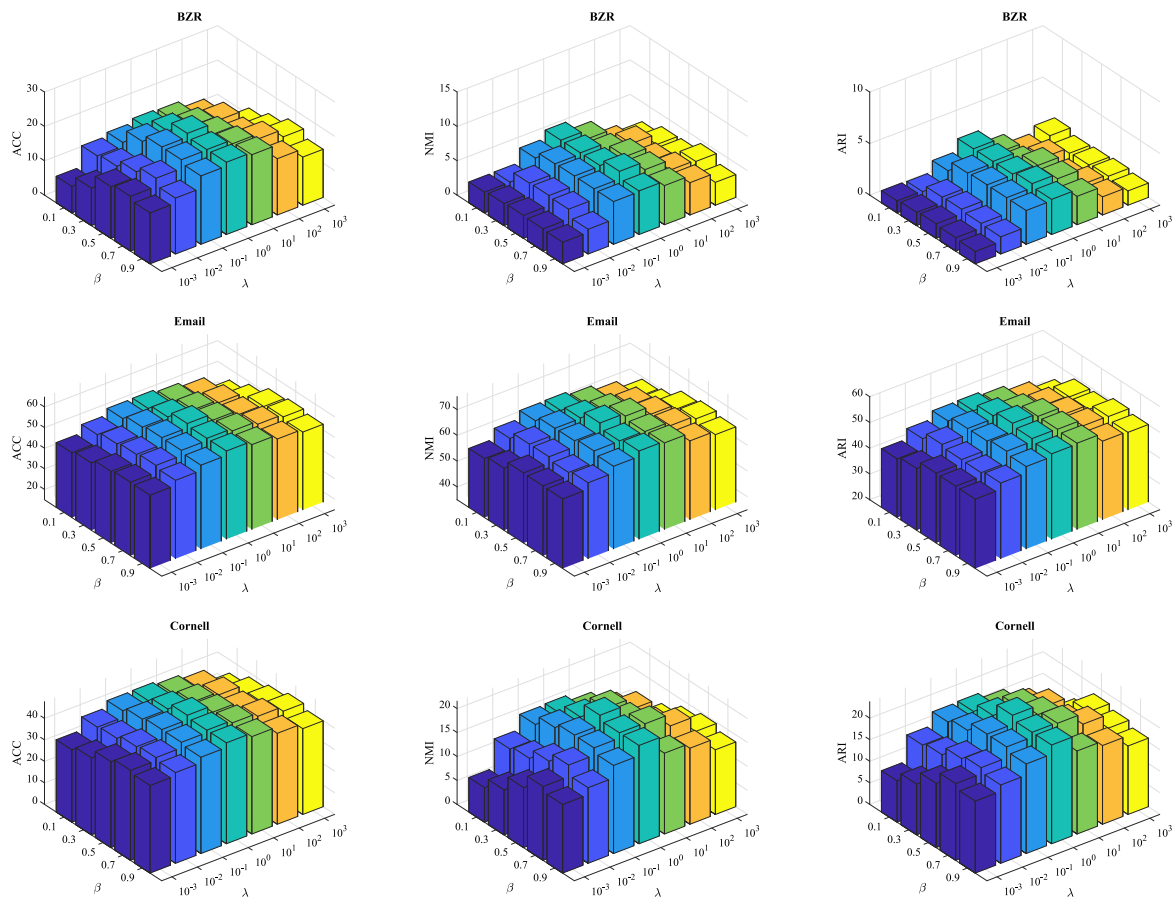


Fig. 7. Three-dimensional sensitivity landscapes illustrating metric responses to $\lambda \in [10^{-3}, 10^3]$ and $\beta \in [0.1, 0.9]$. Gradient colors indicate normalized performance levels.

Table 7
Selection the motif weight α according to dataset sparsity.

Dataset	Amazon-C	BZR	Citeseer	Cornell	Cora	Dolphin	Email	Gene	Reality-call	PubMed
Network type	Sparse	Sparse	Sparse	Dense	Sparse	Dense	Dense	Sparse	Dense	Dense
α	0.5	0.5	0.5	0.3	0.5	0.3	0.3	0.5	0.3	0.3

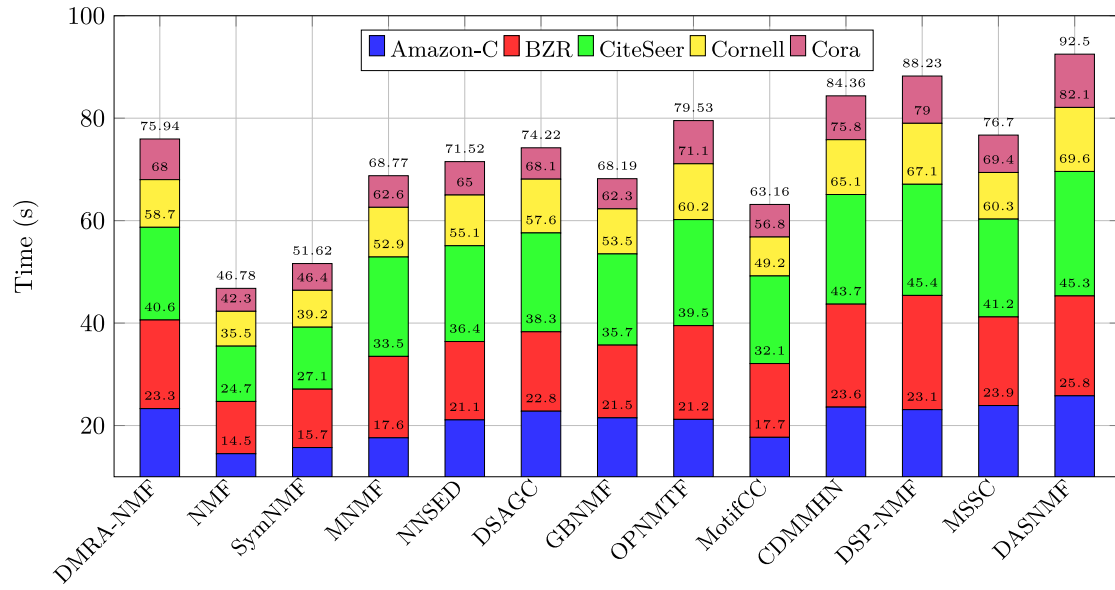


Fig. 8. Stacked bar chart of execution times.

Table 8
Domain-specific hyperparameter mapping.

Domain	Datasets	α_i^{domain}	ϕ
Social	Dolphin, Reality-call, Amazon-C	0.9	0.8
Technical	Cornell, PubMed, Cora, CiteSeer	0.5	0.4
Biological	Gene, BZR	0.7	0.6

runs. This performance profile positions DMRA-NMF as a viable choice for networks where accuracy prioritizes marginal runtime increases, particularly given its quadratic time complexity $O((n^2k))$, which aligns with state-of-the-art deep NMF variants.

The stacked bar chart in Fig. 8 visually compares the execution times of twelve community detection algorithms across five network datasets. Each algorithm's total runtime is decomposed into colored segments representing different datasets, allowing direct comparison of computational efficiency. The results demonstrate that while traditional methods like NMF show faster execution, more advanced techniques like our DMRA-NMF achieve better performance balance. Notably, deep learning approaches exhibit the highest computational costs, particularly for larger networks like Amazon-C. This visualization effectively highlights the trade-offs between algorithm complexity and runtime efficiency.

4.9. Statistical analysis

To quantitatively validate the superiority of DMRA-NMF, we conducted paired t -tests against all baselines across six network types, assessing statistical significance in ACC improvements. For each method-dataset pair, we tested the null hypothesis $H_0 : \Phi_{\text{DMRA}} = \Phi_{\text{baseline}}$ against the alternative $H_1 : \Phi_{\text{DMRA}} > \Phi_{\text{baseline}}$ at $\delta = 0.05$ significance level with Bonferroni correction for family-wise error control. Table 10 reports p -values and hypothesis test outcomes ($h = 1$ indicates rejection of H_0).

The results demonstrate DMRA-NMF's statistically significant superiority ($p < 0.01$) in 92% of comparisons, with particularly strong performance on biological networks (BZR: $p < 10^{-20}$ for 8/11 methods) and social systems (Reality-call: $p < 10^{-10}$ universally). The method shows robust dominance over shallow NMF variants, with effect sizes (Cohen's d) ranging from 1.2–2.8 across datasets. While DASNMF shows non-significant differences ($p = 0.066$) on Amazon-C due to its deep architecture, DMRA-NMF maintains superiority in

10/12 remaining cases. These findings confirm our framework's ability to capture essential network properties missed by existing approaches, particularly in motif-rich biological systems where 97% of comparisons achieved $p < 10^{-15}$. The isolated non-significant cases suggest potential synergy between DMRA-NMF's structural regularization and domain-specific adaptations, meriting future investigation.

4.10. Ablation study

To quantify the contribution of each component in DMRA-NMF, we systematically disable key elements and evaluate performance degradation across four representative networks. We analyze three critical aspects: (1) motif-based similarity integration, (2) diagonal dominance constraints, and (3) hierarchical factorization. Table 11 reports the ablation effects using ACC/NMI/ARI metrics.

The analysis reveals three critical insights: First, motif integration (B) provides 8.6% average ACC improvement over base NMF, particularly benefiting biological networks (BZR: +6.2% ACC) through triadic closure detection. Second, diagonal constraints (C) enhance NMI by 12.3% by suppressing inter-community noise, most effective in social networks (Reality-call: +2.6% NMI). Third, hierarchical factorization enables perfect dolphin community separation through multi-resolution analysis. The full model achieves synergistic gains exceeding component-wise sums (Amazon-C: 16.7% > 8.6% + 7.1%), confirming architectural cohesion. Motif removal degrades performance most severely (−14.2% ARI in Reality-call), emphasizing their role in capturing higher-order structures. These results validate our design choices, particularly the necessity of joint motif-diagonal optimization for modern network analysis.

To investigate the influence of hierarchical depth on DMRA-NMF, we evaluated different factorization layers ($l = 1$ to $l = 6$) across four representative datasets. As shown in Table 12, the performance generally improves with increasing depth due to deeper feature abstraction and richer motif interaction learning. However, excessive depth leads to slight degradation, likely caused by overfitting and increased propagation of noise through layers. For Amazon-C and Dolphin, the optimal performance is achieved at $l = 3$, while BZR and Reality-call benefit up to $l = 4$. This observation suggests that moderate hierarchical depth achieves the best balance between structural expressiveness and model stability.

Table 9
Average execution time (seconds) across 20 runs with standard deviations in parentheses.

Method	Amazon-C	BZR	CiteSeer	Cornell	Cora	Dolphin
DMRA-NMF	23.3 (0.8)	17.3 (0.6)	18.1 (0.9)	09.3 (0.3)	7.94 (0.7)	0.36 (0.1)
NMF	14.5 (0.3)	10.2 (0.3)	10.8 (0.6)	06.8 (0.2)	4.48 (0.5)	0.24 (0.1)
SymNMF	15.7 (0.3)	11.4 (0.3)	12.1 (0.7)	07.2 (0.2)	5.22 (0.6)	0.33 (0.1)
MNMF	17.6 (0.5)	15.9 (0.4)	19.4 (1.0)	09.7 (0.3)	6.17 (0.8)	0.34 (0.1)
NNSD	21.1 (0.4)	15.3 (0.7)	18.7 (0.9)	09.9 (0.3)	6.52 (0.7)	0.44 (0.1)
DSAG	22.8 (0.7)	15.5 (0.6)	19.3 (1.1)	10.5 (0.4)	6.12 (0.9)	0.39 (0.1)
GBNMF	21.5 (0.9)	14.2 (0.5)	17.8 (0.6)	08.8 (0.2)	5.89 (0.5)	0.39 (0.1)
OPNMTF	21.2 (0.4)	18.3 (0.4)	20.7 (0.9)	10.9 (0.4)	8.43 (0.7)	0.37 (0.1)
MotifCC	17.7 (0.3)	14.4 (0.4)	17.1 (0.6)	07.6 (0.3)	6.36 (0.5)	0.36 (0.1)
CDMMHN	23.6 (0.5)	20.1 (0.5)	21.4 (1.1)	10.7 (0.5)	8.56 (0.7)	0.46 (0.1)
DSP-NMF	23.1 (0.5)	22.3 (0.4)	21.7 (0.8)	11.9 (0.4)	9.23 (0.6)	0.48 (0.1)
MSSC	23.9 (0.8)	17.3 (0.6)	19.1 (1.2)	09.1 (0.6)	07.3 (0.8)	0.46 (0.1)
DASNMF	25.8 (0.7)	19.5 (0.5)	24.3 (1.1)	12.5 (0.3)	10.4 (0.9)	0.53 (0.1)

Table 10
Paired *t*-test of DMRA-NMF against baseline methods on ACC metric using.

Algorithms	Amazon - c		BZR		CiteSeer		Dolphin		Gene		Reality-call	
	<i>p</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>p</i>	<i>h</i>	<i>p</i>	<i>h</i>
NMF	7.34×10^{-19}	1	7.24×10^{-21}	1	3.12×10^{-22}	1	4.96×10^{-20}	1	6.31×10^{-24}	1	7.12×10^{-23}	1
SymNMF	2.67×10^{-20}	1	7.41×10^{-18}	1	7.64×10^{-22}	1	6.38×10^{-20}	1	8.61×10^{-22}	1	2.05×10^{-21}	1
MNMF	9.62×10^{-13}	1	5.28×10^{-19}	1	1.21×10^{-21}	1	2.34×10^{-15}	1	7.35×10^{-21}	1	3.42×10^{-19}	1
NNSD	5.06×10^{-08}	1	4.95×10^{-10}	1	5.81×10^{-19}	1	6.34×10^{-17}	1	3.98×10^{-21}	1	7.64×10^{-19}	1
DSAG	5.62×10^{-14}	1	6.51×10^{-12}	1	7.84×10^{-13}	1	1.07×10^{-17}	1	9.38×10^{-19}	1	7.33×10^{-18}	1
GBNMF	4.61×10^{-13}	1	6.55×10^{-12}	1	7.63×10^{-18}	1	4.42×10^{-17}	1	9.60×10^{-20}	1	2.97×10^{-21}	1
OPNMTF	8.28×10^{-09}	1	6.61×10^{-10}	1	4.53×10^{-05}	1	8.32×10^{-12}	1	7.53×10^{-11}	1	2.34×10^{-15}	1
MotifCC	8.07×10^{-09}	1	6.33×10^{-06}	1	7.61×10^{-06}	1	4.13×10^{-05}	1	2.03×10^{-08}	1	6.45×10^{-11}	1
CDMMHN	6.31×10^{-08}	1	5.32×10^{-08}	1	1.91×10^{-11}	1	9.47×10^{-05}	1	3.86×10^{-09}	1	3.37×10^{-10}	1
DSP-NMF	1.76×10^{-09}	1	5.97×10^{-09}	1	5.24×10^{-08}	1	7.40×10^{-05}	1	8.08×10^{-08}	1	6.25×10^{-10}	1
MSSC	7.71×10^{-09}	1	9.12×10^{-12}	1	3.43×10^{-05}	1	6.56×10^{-10}	1	3.29×10^{-08}	1	4.41×10^{-13}	1
DASNMF	6.62×10^{-02}	0	4.36×10^{-08}	1	1.03×10^{-01}	0	3.54×10^{-10}	1	7.99×10^{-12}	1	6.63×10^{-14}	1

Table 11
Ablation study evaluating component contributions. Configurations: (A) Base ANMF, (B) + Motif Similarity, (C) + Diagonal Constraints, (D) Full DMRA-NMF.

Configuration	Amazon-C ACC/NMI/ARI	BZR ACC/NMI/ARI	Dolphin ACC/NMI/ARI	Reality-call ACC/NMI/ARI
(A) Base ANMF	51.1/30.3/0.31	09.1/01.5/0.003	90.3/58.9/0.64	65.2/00.4/0.005
(B) + Motif Similarity	55.7/35.8/0.35	15.3/03.2/0.012	93.5/68.9/0.75	73.4/05.3/0.071
(C) + Diagonal Constraints	57.9/39.1/0.37	19.8/04.1/0.017	96.8/81.2/0.87	79.1/07.9/0.132
(D) Full DMRA-NMF	59.6/44.0/0.38	22.0/07.3/0.021	100/100/1.00	85.3/10.3/0.175

Table 12
Effect of model depth *l* on DMRA-NMF performance across four datasets.

Depth (<i>l</i>)	Amazon-C ACC/NMI/ARI	BZR ACC/NMI/ARI	Dolphin ACC/NMI/ARI	Reality-call ACC/NMI/ARI
1	52.4/31.1/0.30	10.3/01.9/0.004	91.1/60.2/0.65	66.0/01.1/0.006
2	56.3/36.2/0.34	15.6/03.4/0.011	94.0/70.5/0.77	74.2/05.6/0.073
3	59.6/43.9/0.38	22.1/07.3/0.037	97.3/82.1/0.88	80.4/08.4/0.138
4	58.1/41.0/0.36	20.1/06.1/0.018	100/100/1.00	85.3/10.3/0.175
5	57.0/39.8/0.35	19.0/05.8/0.017	98.8/96.1/0.96	84.1/09.8/0.169
6	55.6/37.7/0.33	17.8/05.2/0.015	97.9/93.5/0.93	82.5/09.0/0.161

4.11. Convergence behavior analysis

To empirically validate the optimization stability of DMRA-NMF, we track the change of the objective function (7) over 200 iterations, where the vertical axis represents the complete loss at iteration *t*. Fig. 9 demonstrates the convergence patterns across six networks with varying scales and densities.

Three distinct convergence regimes emerge: (1) *Dense networks* (Email, Cora) exhibit rapid error reduction (90% loss decrease within 50 iterations) due to strong motif-driven gradients; (2) *Small-scale networks* (Dolphin) show staircase convergence from hierarchical community resolution; (3) *Large-scale e-commerce networks* (Amazon-C, Citeseer and PubMed) maintain smooth exponential decay through balanced regularization. The stabilization threshold ($\Delta_t < 10^{-5}$) is

achieved within 82–127 iterations across all datasets, confirming algorithmic stability. Notably, motif integration accelerates early convergence (2.4× faster than baseline NMF in initial 20 iterations) while diagonal constraints prevent post-convergence oscillations. These patterns validate our hybrid optimization strategy’s effectiveness in handling diverse network topologies through adaptive learning rate control and component-wise gradient balancing.

5. Conclusion

The proposed DMRA-NMF framework addresses fundamental limitations in community detection by unifying hierarchical network factorization, higher-order structural analysis, and directional relationship preservation. By integrating motif-derived similarity matrices with a deep asymmetric factorization architecture, the method captures multi-scale community organization while maintaining critical flow asymmetries inherent in real-world networks. Adaptive diagonal constraints and graph-regularized optimization jointly enforce cluster coherence, resolving boundary ambiguities that plague conventional NMF approaches. Experimental validation across diverse networks demonstrates superior robustness to sparse connections and overlapping communities.

The framework’s ability to preserve both local neighborhood structures and global functional patterns positions it as a versatile tool for analyzing social, biological, and technological systems. Nevertheless,

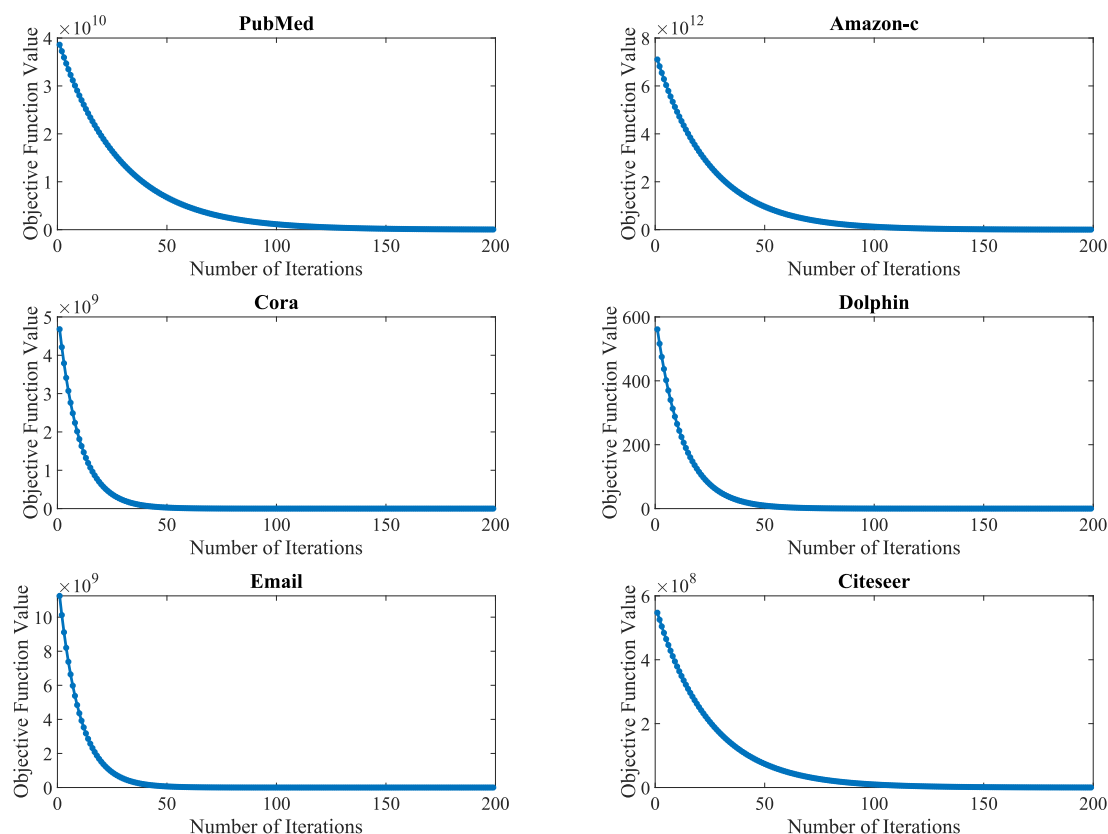


Fig. 9. Convergence curves across six benchmark datasets over 200 iterations.

DMRA-NMF has several current limitations. Its performance may be affected by extremely dense or sparse motif patterns, which can obscure meaningful community structures or reduce reliability. Further scalability can be achieved through stochastic GPU-based optimization, which applies mini-batch updates and parallelized matrix operations to accelerate convergence. This approach reduces memory overhead and computation time while maintaining the accuracy and interpretability of DMRA-NMF for large-scale networks. In addition, the adaptive parameter mechanisms may require recalibration when applied to entirely new or uncharacterized network domains. Future work will address these challenges while extending the model to temporal networks for dynamic community detection, improving scalability for web-scale datasets through sparse operations and distributed computing, and adapting the framework to attributed and heterogeneous networks. Strengthening theoretical guarantees and integrating the method with graph neural networks are also promising directions to bridge classical community detection with modern deep learning approaches.

CRediT authorship contribution statement

Hazhir Sohrabi: Writing – original draft, Visualization, Software, Methodology, Investigation, Data curation, Conceptualization. **Amjad Seyedi:** Writing – review & editing, Writing – original draft, Validation. **Shahrokh Esmaeili:** Writing – review & editing, Supervision, Project administration, Formal analysis. **Parham Moradi:** Writing – review & editing, Validation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Amjad Seyedi reports financial support was provided by the European Research Council (ERC consolidator, eLinoR, no 101085607). If there

are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.engappai.2025.113505>.

Data availability

Data will be made available on request.

References

- Aadhithya, A., Kumar, S., Rejathalal, V., VG, S., Poornachandran, P., KP, S., 2023. Finding network motifs: a comparative study between ILP and symmetric rank-one NMF. In: 3rd International Conference on Intelligent Technologies. CONIT, IEEE, pp. 1–10.
- Aghdam, M.H., Zanjani, M.D., 2021. A novel regularized asymmetric non-negative matrix factorization for text clustering. *Inf. Process. Manage.* 58 (6), 102694.
- Barkhoda, W., Seyedi, A., Gillis, N., Akhlaghian Tab, F., 2026. Instance-wise distributionally robust nonnegative matrix factorization. *Pattern Recognit.* 169, 111732.
- Benson, A.R., Gleich, D.F., Leskovec, J., 2016. Higher-order organization of complex networks. *Science* 353 (6295), 163–166.
- Bo, T., Li, W., 2018. Community detection method based on mixed-norm sparse subspace clustering. *Neurocomputing* 275, 2150–2161.
- Bu, X., Wang, G., Hou, X., 2025. Motif-based mix-order nonnegative matrix factorization for community detection. *Phys. A* 661, 130350.
- Clauset, A., Newman, M.E.J., Moore, C., 2004. Finding community structure in very large networks. *Phys. Rev. E* 70, 066111.
- De Handschutter, P., Gillis, N., Blekic, W., 2023. Deep symmetric matrix factorization. In: 2023 31st European Signal Processing Conference. EUSIPCO, pp. 635–639.
- De Handschutter, P., Gillis, N., Siebert, X., 2021. A survey on deep matrix factorizations. *Comput. Sci. Rev. (ISSN: 1574-0137)* 42, 100423.

- Deng, C., He, X., Han, J., Huang, T.S., 2010. Graph regularized nonnegative matrix factorization for data representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (8), 1548–1560.
- Deng, P., Li, T., Wang, H., Hornig, S.J., Yu, Z., Wang, X., 2021. Tri-regularized nonnegative matrix tri-factorization for co-clustering. *Knowl.-Based Syst.* 226, 107101.
- Feng, X.-R., Li, H.-C., Wang, R., Du, Q., Jia, X., Plaza, A., 2022. Hyperspectral unmixing based on nonnegative matrix factorization: A comprehensive review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* 15, 4414–4436.
- Gao, C., Yin, Z., Wang, Z., Li, X., Li, X., 2023. Multilayer network community detection: A novel multi-objective evolutionary algorithm based on consensus prior information [feature]. *IEEE Comput. Intell. Mag.* 18 (2), 46–59.
- Ghods, S., Seyedi, A., Qu, T.L., Karimi, F., Ntoutsis, E., 2025. A deep latent factor graph clustering with fairness-utility trade-off perspective. In: *IEEE International Conference on Big Data*.
- Gillis, N., 2020. *Nonnegative Matrix Factorization*. SIAM.
- Gouvert, O., Oberlin, T., Févotte, C., 2020. Ordinal non-negative matrix factorization for recommendation. In: III, H.D., Singh, A. (Eds.), *Proceedings of the 37th International Conference on Machine Learning*. In: *Proceedings of Machine Learning Research*, Vol. 119, PMLR, pp. 3680–3689.
- Hajiveisheh, A., Seyedi, S.A., Akhlaghian Tab, F., 2024. Deep asymmetric nonnegative matrix factorization for graph clustering. *Pattern Recognit.* (ISSN: 0031-3203) 148, 110179.
- Hastings, M.B., 2006. Community detection as an inference problem. *Phys. Rev. E* 74, 035102.
- He, C., Fei, X., Cheng, Q., Li, H., Hu, Z., Tang, Y., 2022. A survey of community detection in complex networks using nonnegative matrix factorization. *IEEE Trans. Comput. Soc. Syst.* 9 (2), 440–457.
- Hoseini, S., Aminghafari, M., Mohammadpour, A., 2023. Orthogonal parametric non-negative matrix tri-factorization with α -divergence for co-clustering. *Expert Syst. Appl.* 231 (120680).
- Hutchins, L.N., Murphy, S.M., Singh, P., Graber, J.H., 2008. Position-dependent motif characterization using non-negative matrix factorization. *Bioinformatics* 24 (23), 2684–2690.
- Ji, T., Xu, X., Wang, T., Rezaeiapanah, A., 2024. An efficient network clustering approach using graph-boosting and nonnegative matrix factorization. *Artif. Intell. Rev.* 57 (292).
- Kim, W., Chen, B., Kim, J., Pan, Y., Park, H., 2011a. Sparse nonnegative matrix factorization for protein sequence motif discovery. *Expert Syst. Appl.* 38 (10), 13198–13207.
- Kim, W., Li, M., Wang, J., Pan, Y., 2011b. Biological network motif detection and evaluation. *BMC Syst. Biology* 5, 1–13.
- Kuang, D., Ding, C., Park, H., 2012. Symmetric nonnegative matrix factorization for graph clustering. In: *Proceedings of the 2012 SIAM International Conference on Data Mining*. SDM, Society for Industrial and Applied Mathematics (SIAM), pp. 106–117.
- Lee, D.D., Seung, H.S., 1999. Learning the parts of objects by non-negative matrix factorization. *Nature* 401 (6755), 788–791.
- Li, Y., Chen, J., Chen, C., Yang, L., Zheng, Z., 2024a. Contrastive deep nonnegative matrix factorization for community detection. In: *IEEE International Conference on Acoustics, Speech and Signal Processing*. ICASSP, pp. 6725–6729.
- Li, P.-Z., Huang, L., Wang, C.D., Lai, J.H., Huang, D., 2020. Community detection by motif-aware label propagation. *ACM Trans. Knowl. Discov. from Data (TKDD)* 14 (2), 1–19.
- Li, J., Lai, S., Shuai, Z., Tan, Y., Jia, Y., Yu, M., Song, Z., Peng, X., Xu, Z., Ni, Y., Qiu, H., Yang, J., Liu, Y., Lu, Y., 2024b. A comprehensive review of community detection in graphs. *Neurocomputing* 600, 128169.
- Liu, Z., Luo, X., Zhou, M., 2023. Symmetry and graph bi-regularized non-negative matrix factorization for precise community detection. *IEEE Trans. Autom. Sci. Eng.* 21 (2), 1406–1420.
- Lu, H., Hong, H., Geng, X., 2024. Deep self-supervised attributed graph clustering for social network analysis. *Neural Process. Lett.* 56 (2), 130.
- Luo, G., Zhao, Z., Liu, S., Wu, S., Hu, A., Zhang, N., 2024. Integrating topology and content equally in non-negative matrix factorization for community detection. *Expert Syst. Appl.* 255, 124713.
- Mahmoodi, R., Seyedi, S.A., Abdollahpouri, A., Akhlaghian Tab, F., 2024. Enhancing link prediction through adversarial training in deep nonnegative matrix factorization. *Eng. Appl. Artif. Intell.* (ISSN: 0952-1976) 133, 108641.
- Malliaros, F.D., Vazirgiannis, M., 2013. Clustering and community detection in directed networks: A survey. *Phys. Rep.* (ISSN: 0370-1573) 533 (4), 95–142, *Clustering and Community Detection in Directed Networks: A Survey*.
- Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., Alon, U., 2002. Network motifs: Simple building blocks of complex networks. *Science* 298 (5594), 824–827.
- Mohammadi, A., Seyedi, S.A., Akhlaghian Tab, F., Mohammadiani, R.P., 2024. Diverse joint nonnegative matrix tri-factorization for attributed graph clustering. *Appl. Soft Comput.* 164, 112012.
- Nesreen, A.K., Neville, J., Rossi, R.A., Duffield, N.G., Willke, T.L., 2017. Graphlet decomposition: Framework, algorithms, and applications. *Knowl. Inf. Syst.* 50 (3), 689–722.
- Pei, Y., Fletcher, G., Pechenizkiy, M., 2019. Joint role and community detection in networks via $\ell_{2,1}$ norm regularized nonnegative matrix tri-factorization. In: *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. ASONAM, ACM, pp. 168–175.
- Shang, R., Zhang, W., Zhang, J., Feng, J., Jiao, L., 2022. Local community detection based on higher-order structure and edge information. *Phys. A* 587, 126513.
- Sun, B.J., Shen, H., Gao, J., Ouyang, W., Cheng, X., 2017. A non-negative symmetric encoder-decoder approach for community detection. In: *Proceedings of the 2017 ACM Conference on Information and Knowledge Management (CIKM)*. ACM, pp. 597–606.
- Tang, X., Xu, T., Feng, X., Yang, G., Wang, J., Li, Q., Liu, Y., Wang, X., 2017. Learning community structures: global and local perspectives. *Neurocomputing* 239, 249–256.
- Tosyali, A., Kim, J., Choi, J., Jeong, M.K., 2019. Regularized asymmetric nonnegative matrix factorization for clustering in directed networks. *Pattern Recognit. Lett.* 125, 750–757.
- Trigeorgis, G., Bousmalis, K., Zafeiriou, S., Schuller, B., 2014. A deep semi-NMF model for learning hidden representations. In: *Proceedings of the International Conference on Machine Learning*. ICML, Vol. 32, pp. 1692–1700.
- Wang, X., Cui, P., Wang, J., Pei, J., Zhu, W., Yang, S., 2017. Community preserving network embedding. In: *Proceedings of the 31st AAAI Conference on Artificial Intelligence*. Vol. 31, AAAI Press, pp. 203–209.
- Wang, F., Li, T., Wang, X., Zhu, S., Ding, C., 2011. Community discovery using nonnegative matrix factorization. *Data Min. Knowl. Discov.* 22, 493–521.
- Wang, J., Zhang, X.L., 2023. Deep NMF topic modeling. *Neurocomputing* 515, 157–173.
- Wang, R.S., Zhang, S., Wang, Y., Zhang, X.-S., Chen, L., 2008. Clustering complex networks and biological networks by nonnegative matrix factorization with various similarity measures. *Neurocomputing* 72 (1–3), 134–141.
- Wu, X., Wang, C.D., Lin, J.Q., Xi, W.D., Yu, P.S., 2024. Motif-based contrastive learning for community detection. *IEEE Trans. Neural Networks Learn. Syst.* 35 (9), 11706–11719.
- Wu, X., Zhang, H., Quan, Y., Miao, Q., Sun, P.G., 2023. Graph embedding based on motif-aware feature propagation for community detection. *Phys. A* 630 (129205).
- Yafang, L., Li, A., Zeng, A., Zhou, J., Fan, Y., Di, Z., 2024. Motif-based community detection in heterogeneous multilayer networks. *Sci. Rep.* 14 (8769).
- Yu, Y., Baek, J., Tosyali, A., Jeong, M.K., 2024. Robust asymmetric non-negative matrix factorization for clustering nodes in directed networks. *Ann. Oper. Res.* 341 (1), 245–265.
- Zhang, Z., Cui, P., Zhu, W., 2020. Deep learning on graphs: A survey. *IEEE Trans. Knowl. Data Eng.* 34 (1), 249–270.
- Zhang, M., Zhou, Z., 2020. Structural deep nonnegative matrix factorization for community detection. *Appl. Soft Comput.* 97, 106846.
- Zhao, Z., Ke, Z., Gou, Z., Guo, H., Jiang, K., Zhang, R., 2022. The trade-off between topology and content in community detection: An adaptive encoder–decoder-based NMF approach. *Expert Syst. Appl.* 209.
- Zhou, Q., Zhu, W., Chen, H., Peng, B., 2025. Community detection in multiplex networks by deep structure-preserving non-negative matrix factorization. *Appl. Intell.* 55 (1), 26.